

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 13:35:22 ; Search time 88.4672 Seconds
(without alignments)
1194.261 Million cell updates/sec

Title: US-10-037-860-4

Perfect score: 1729

Sequence: 1 MAMTLLDWCRCMDVNSQRT.....LTGAGEGPKPLSVAGADP 329

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1342398 seqs, 321133274 residues

Total number of hits satisfying chosen parameters: 1342398

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	ID	Description
1	1729	100.0	329	13 US-10-037-860-4	Sequence 4, Appli
2	1666	96.4	318	12 US-09-804-014A-40	Sequence 40, Appl
3	1602	92.7	353	9 US-09-965-529-7	Sequence 7, Appli
4	1602	92.7	353	10 US-09-969-680A-7	Sequence 7, Appli
5	887.5	51.3	321	12 US-09-804-014A-39	Sequence 39, Appl
6	887.5	51.3	351	9 US-09-965-529-1	Sequence 1, Appli
7	887.5	51.3	351	10 US-09-969-680A-1	Sequence 1, Appli
8	887.5	51.3	351	12 US-09-804-014A-16	Sequence 16, Appl
9	887.5	51.3	351	15 US-10-341-434-10	Sequence 10, Appl
10	874.5	50.6	312	12 US-09-804-014A-73	Sequence 73, Appl
11	874.5	50.6	312	12 US-09-804-014A-74	Sequence 74, Appl
12	766.5	44.3	463	13 US-10-037-860-13	Sequence 13, Appl
13	744	43.0	452	16 US-10-408-765A-2385	Sequence 2385, Ap
14	620	35.9	299	15 US-10-094-749-1978	Sequence 1978, Ap
15	564	32.6	283	13 US-10-037-860-11	Sequence 11, Appl

16 441 25.5 195 13 US-10-037-860-7 Sequence 7, Appli
17 353.5 20.4 120 12 US-09-804-014A-41 Sequence 41, Appl
18 338.5 19.6 120 12 US-09-804-014A-42 Sequence 42, Appl
19 335.5 19.4 403 15 US-10-094-466-38 Sequence 38, Appl
20 326 18.9 337 12 US-10-296-115-1208 Sequence 1208, Ap
21 304 17.6 204 14 US-10-029-386-3747 Sequence 3747, A
22 256.5 14.8 149 13 US-10-037-860-9 Sequence 9, Appli
23 246.5 14.3 116 9 US-09-864-761-34645 Sequence 34645, A
24 192 11.1 538 16 US-10-408-765A-2992 Sequence 2992, A
25 133 7.7 584 12 US-10-221-278-355 Sequence 355, App
26 133 7.7 584 15 US-10-291-172-355 Sequence 355, App
27 120.5 7.0 5245 14 US-10-329-079-45 Sequence 45, Appl
28 115.5 6.7 555 12 US-10-282-122A-49641 Sequence 49641, A
29 114.5 6.6 555 12 US-10-282-122A-50770 Sequence 50770, A
30 114 6.6 558 12 US-10-282-122A-65151 Sequence 65151, A
31 111 6.4 558 12 US-10-282-122A-66072 Sequence 66072, A
32 109 6.3 531 15 US-10-369-493-17979 Sequence 17979, A
33 107.5 6.2 526 12 US-10-282-122A-47973 Sequence 47973, A
34 106.5 6.2 503 9 US-09-738-626-5485 Sequence 5485, Ap
35 104.5 6.0 556 12 US-10-282-122A-58461 Sequence 58461, A
36 103 6.0 935 14 US-10-080-608A-25 Sequence 25, Appl
37 103 6.0 935 15 US-10-370-685-114 Sequence 114, App
38 103 6.0 1855 16 US-10-437-963-144559 Sequence 144559, A
39 101.5 5.9 495 12 US-10-425-114-50904 Sequence 50904, A
40 99.5 5.8 1997 16 US-10-408-765A-1013 Sequence 1013, Ap
41 99.5 5.8 5245 14 US-10-329-079-11 Sequence 11, Appl
42 98 5.7 553 12 US-10-282-122A-62912 Sequence 62912, A
43 98 5.7 1009 16 US-10-607-631-2 Sequence 2, Appli
44 98 5.7 1276 14 US-10-156-761-10509 Sequence 10509, A
45 97 5.6 538 12 US-10-282-122A-51122 Sequence 51122, A

ALIGNMENTS

RESULT 1
US-10-037-860-4
; Sequence 4, Application US/10037860
; Publication No. US20020123114A1
; GENERAL INFORMATION:
; APPLICANT: Jerome B. Posner
; APPLICANT: Joseph O. Dalmau
; APPLICANT: Myrna R. Rosenfeld
; TITLE OF INVENTION: MA FAMILY POLYPEPTIDES AND ANTI-MA
; FILE REFERENCE: 2581.1004-004
; CURRENT APPLICATION NUMBER: US/10/037,860
; CURRENT FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 09/189,527
; PRIOR FILING DATE: 1998-11-10
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 329
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-037-860-4

Query Match 100.0%; Score 1729; DB 13; Length 329;
Best Local Similarity 100.0%; Pred. No. 5e166;
Matches 329; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAMTLLDWCRCMDVNSQRTLLVWGPVNCDAEIEETLQAAMPQVSYRMLGRMFWRN 60
Db 1 MAMTLLDWCRCMDVNSQRTLLVWGPVNCDAEIEETLQAAMPQVSYRMLGRMFWRN 60
QY 61 AKAAELLELTGAVDYAAIPREMFGKGVKWLFPKPTSDAEFLERLHLFLAREGWTVDVA 120
Db 61 AKAAELLELTGAVDYAAIPREMFGKGVKWLFPKPTSDAEFLERLHLFLAREGWTVDVA 120
QY 121 RVLGFQNPFTTGPMPAEMLNIIIDNVIOPLVESIWKYKRLTLFSGKGHPRAWRFDPW 180
Db 121 RVLGFQNPFTTGPMPAEMLNIIIDNVIOPLVESIWKYKRLTLFSGKGHPRAWRFDPW 180

Db 181 LEHTNEVLEWQVSDVEKRRRLMESLRGPAADVIRILKSNPAITTAECLEKALQVFGSV 240
QY 241 ESSRDAQIKELNTYQNGEKLISAYVIRLEPLLQKVVEKGAIDKDNVNQARLEQVIAGANH 300
Db 241 ESSRDAQIKELNTYQNGEKLISAYVIRLEPLLQKVVEKGAIDKDNVNQARLEQVIAGANH 300
QY 301 SGAIRROLWLTGAGEGPG 318
Db 301 SGAIRROLWLTGAGEGPG 318
RESULT 3
US-09-965-529-7
; Sequence 7, Application US/09965529
; Publication No. US20020182671A1
; GENERAL INFORMATION:
; APPLICANT: LAL, Preeti
; APPLICANT: YUE, Henry
; APPLICANT: TANG, Y. Tom
; APPLICANT: BANDMAN, Olga
; APPLICANT: BURFORD, Neil
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: BAUGHN, Maria R.
; APPLICANT: LU, Dyung Aina M.
; APPLICANT: PATTERSON, Chandra
; TITLE OF INVENTION: MEMBRANE ASSOCIATED PROTEINS
; FILE REFERENCE: PF-0731 USA
; CURRENT APPLICATION NUMBER: US/09/965,529
; CURRENT FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/149,641; 60/164,203; PCT/US00/22315
; PRIOR FILING DATE: 1999-08-17; 1999-11-09; 2000-08-14
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PERL Program
; SEQ ID NO 7
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020182671A1 2483172CD1
US-09-965-529-7
Query Match 92.7%; Score 1602; DB 9; Length 353;
Best Local Similarity 96.6%; Pred. No. 3.9e-153; Indels 0; Gaps 0;
Matches 308; Conservative 1; Mismatches 10;
QY 1 MAMTLLDWCRCMDVNSQRTLLVWGIPVNCDEAEIEETLQAAAMPQVSYRMLGRMFREEN 60
Db 1 MAMTLLDWCRCMDVNSQRTLLVWGIPVNCDEAEIEETLQAAAMPQVSYRMLGRMFREEN 60
QY 61 AKAAELLELTGAVDYAAIPREMPKGGVWVLFKPTSDAEFLERLHLFLAREGWTVDVA 120
Db 61 AKAAELLELTGAVDYAAIPREMPKGGVWVLFKPTSDAEFLERLHLFLAREGWTVDVA 120
QY 121 RVLGFGQNPPTPGPEMPAEMNLNVIQPLVESIWKRLTLFSGKGHPRAWGNFDPW 180
Db 121 RVLGFGQNPPTPGPEMPAEMNLNVIQPLVESIWKRLTLFSGKGHPRAWGNFDPW 180
QY 181 LEHTNEVLEWQVSDVEKRRRLMESLRGPAADVIRILKSNPAITTAECLEKALQVFGSV 240
Db 181 LEHTNEVLEWQVSDVEKRRRLMESLRGPAADVIRILKSNPAITTAECLEKALQVFGSV 240
QY 241 ESSRDAQIKELNTYQNGEKLISAYVIRLEPLLQKVVEKGAIDKDNVNQARLEQVIAGANH 300
Db 241 ESSRDAQIKELNTYQNGEKLISAYVIRLEPLLQKVVEKGAIDKDNVNQARLEQVIAGANH 300
QY 301 SGAIRROLWLTGAGEGPG 319
Db 301 SGAIRROLWLTGAGEGPG 319
RESULT 4
US-09-969-680A-7

QY 181 LEHTNEVLEWQVSDVEKRRRLMESLRGPAADVIRILKSNPAITTAECLEKALQVFGSV 240
Db 181 LEHTNEVLEWQVSDVEKRRRLMESLRGPAADVIRILKSNPAITTAECLEKALQVFGSV 240
QY 241 ESSRDAQIKELNTYQNGEKLISAYVIRLEPLLQKVVEKGAIDKDNVNQARLEQVIAGANH 300
Db 241 ESSRDAQIKELNTYQNGEKLISAYVIRLEPLLQKVVEKGAIDKDNVNQARLEQVIAGANH 300
QY 301 SGAIRROLWLTGAGEGPGPKPLSVAGADP 329
Db 301 SGAIRROLWLTGAGEGPGPKPLSVAGADP 329
RESULT 2
US-09-804-014A-40
; Sequence 40, Application US/09804014A
; Publication No. US20030064489A1
; GENERAL INFORMATION:
; APPLICANT: Li, Li
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Vernet, Corine
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard
; APPLICANT: Spaderna, Steven
; APPLICANT: Majumder, Kumud
; TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-721 US
; CURRENT APPLICATION NUMBER: US/09/804,014A
; CURRENT FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: 60/188,316
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 60/188,277
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 60/189,139
; PRIOR FILING DATE: 2000-03-14
; PRIOR APPLICATION NUMBER: 60/189,140
; PRIOR FILING DATE: 2000-03-14
; PRIOR APPLICATION NUMBER: 60/190,401
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/190,231
; PRIOR FILING DATE: 2000-03-17
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 40
; LENGTH: 318
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: VARIANT
; LOCATION: (20)
; OTHER INFORMATION: Wherein Xaa is any amino acid as defined in the specification
; OTHER INFORMATION: specification
US-09-804-014A-40
Query Match 96.4%; Score 1666; DB 12; Length 318;
Best Local Similarity 99.7%; Pred. No. 1.1e-159; Indels 0; Gaps 0;
Matches 317; Conservative 0; Mismatches 1;
QY 1 MAMTLLDWCRCMDVNSQRTLLVWGIPVNCDEAEIEETLQAAAMPQVSYRMLGRMFREEN 60
Db 1 MAMTLLDWCRCMDVNSQRTLLVWGIPVNCDEAEIEETLQAAAMPQVSYRMLGRMFREEN 60
QY 61 AKAAELLELTGAVDYAAIPREMPKGGVWVLFKPTSDAEFLERLHLFLAREGWTVDVA 120
Db 61 AKAAELLELTGAVDYAAIPREMPKGGVWVLFKPTSDAEFLERLHLFLAREGWTVDVA 120
QY 121 RVLGFGQNPPTPGPEMPAEMNLNVIQPLVESIWKRLTLFSGKGHPRAWGNFDPW 180
Db 121 RVLGFGQNPPTPGPEMPAEMNLNVIQPLVESIWKRLTLFSGKGHPRAWGNFDPW 180
QY 181 LEHTNEVLEWQVSDVEKRRRLMESLRGPAADVIRILKSNPAITTAECLEKALQVFGSV 240

```
; Sequence 7, Application US/09969680A
; Publication No. US20030124649A1
; GENERAL INFORMATION:
; APPLICANT: LAL, Preeti; YUE, Henry
; APPLICANT: TANG, Y. Tom; BANDMAN, Olga
; APPLICANT: BURFORD, Neil; AZIMZAI, Yalda
; APPLICANT: BAUGHN, Mariah R.; LU, Dyung Aina M.
; APPLICANT: PATTERSON, Chandra
; TITLE OF INVENTION: MEMBRANE ASSOCIATED PROTEINS
; FILE REFERENCE: PF-0731-1 USA
; CURRENT APPLICATION NUMBER: US/09/969,680A
; CURRENT FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US00/22315
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/149,641
; PRIOR FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: 60/164,203
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PERL Program
; SEQ ID NO 7
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. US20030124649A1 2483172CD1
US-09-969-680A-7

Query Match          92.7%; Score 1602; DB 10; Length 353;
Best Local Similarity 96.6%; Pred. No. 3.9e-153;
Matches 308; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

QY 1 MAMTLLDWCRCMDVNSQRTLLVWGIPVNCDEABIEETLQAAIPQVSVYMLGRMFWEEN 60
Db 1 MAMTLLDWCRCMDVNSQRTLLVWGIPVNCDEABIEETLQAAIPQVSVYMLGRMFWEEN 60
QY 61 AKAAELLELTGAVDYAAIPREMPGKGGVWVLFKPPPTSDAEFLERLHLFLAREGWTVDVA 120
Db 61 AKAAELLELTGAVDYAAIPREMPGKGGVWVLFKPPPTSDAEFLERLHLFLAREGWTVDVA 120
QY 121 RVLGFGNPTTPGPEPAEMNLVLDNVIQPLVESIWKRLTLFSGKGHPRAWRG 180
Db 121 RVLGFGNPTTPGPEPAEMNLVLDNVIQPLVESIWKRLTLFSGKGHPRAWRG 180
QY 181 LEHNEVLEEWQVSDVEKRRRLMESLRGPAADVIRILKSNPAITTAECLEKALEQ 240
Db 181 LEHNEVLEEWQVSDVEKRRRLMESLRGPAADVIRILKSNPAITTAECLEKALEQ 240
QY 241 ESSRDQIKPLNTYQNPGEKLSAYVIRLEPPLQKVEKGAKDKNNVQARLEQVIAGANH 300
Db 241 ESSRDQIKPLNTYQNPGEKLSAYVIRLEPPLQKVEKGAKDKNNVQARLEQVIAGANH 300
QY 301 SGAIRQLMLTGAGEGPG 319
Db 301 SGAIRQLMLTGAGEGPG 319

RESULT 5
US-09-804-014A-39
; Sequence 39, Application US/09804014A
; Publication No. US20030064489A1
; GENERAL INFORMATION:
; APPLICANT: Li, Li
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Vernet, Corine
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard
; APPLICANT: Spaderna, Steven
; APPLICANT: Majumder, Kumud
; TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-721 US
; CURRENT APPLICATION NUMBER: US/09/804,014A
```

```
; CURRENT FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: 60/188,316
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 60/188,277
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 60/189,139
; PRIOR FILING DATE: 2000-03-14
; PRIOR APPLICATION NUMBER: 60/189,140
; PRIOR FILING DATE: 2000-03-14
; PRIOR APPLICATION NUMBER: 60/190,401
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/190,231
; PRIOR FILING DATE: 2000-03-17
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 39
; LENGTH: 321
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-804-014A-39

Query Match          51.3%; Score 887.5; DB 12; Length 321;
Best Local Similarity 55.4%; Pred. No. 7.1e-81;
Matches 179; Conservative 53; Mismatches 84; Indels 7; Gaps 5;

QY 1 MAMTLLDWCRCMDVNSQRTLLVWGIPVNCDEABIEETLQAAIPQVSVYMLGRMFWEEN 59
Db 1 MTELLDWCRCMDVNSQRTLLVWGIPVNCDEABIEETLQAAIPQVSVYMLGRMFWEEN 60
QY 60 NAKAAELLELTGAVDYAAIPREMPGKGGVWVLFKPPPTSDAEFLERLHLFLAREGWTVDV 119
Db 61 NRKVALVGLTAETSHALVPEIKPGGIGWVIFKPPDPDNTFSLRNEFLAGEGMYGEL 120
QY 120 ARVLGFQNPPT--PTPG--PEMPAEMNLVLDNVIQPLVESIWKRLTLFSGKGHPRAWRG 175
Db 121 SRALGHENGSLDPEQGMIPENWAPMLAQALE-ALQPALQCLKYKLVFSGRESPEFGE 179
QY 176 NFDPELHTNEVLEEWQVSDVEKRRRLMESLRGPAADVIRILKSNPAITTAECLEKALEQ 235
Db 180 EFGRWPFHTTQMIAKAWQVDPVEKRRRLMESLRGPAADVIRILKSNPAITTAECLEKALEE 239
QY 236 VFGSVSSRDQIKPLNTYQNPGEKLSAYVIRLEPPLQKVEKGAKDKNNVQARLEQVI 295
Db 240 VFGVTNPRLQKYLITTYQKDEEKLSAYVIRLEPPLQKVEKGAKDKNNVQARLEQVI 299
QY 296 AGAHSGAIRQLMLTGAGEGPG 318
Db 300 AGAVHK-TIRRELNPEDGPGAPG 321

RESULT 6
US-09-965-529-1
; Sequence 1, Application US/09965529
; Publication No. US20020182671A1
; GENERAL INFORMATION:
; APPLICANT: LAL, Preeti
; APPLICANT: YUE, Henry
; APPLICANT: TANG, Y. Tom
; APPLICANT: BANDMAN, Olga
; APPLICANT: BURFORD, Neil
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: LU, Dyung Aina M.
; APPLICANT: PATTERSON, Chandra
; TITLE OF INVENTION: MEMBRANE ASSOCIATED PROTEINS
; FILE REFERENCE: PF-0731 USA
; CURRENT APPLICATION NUMBER: US/09/965,529
; CURRENT FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/149,641; 60/164,203; PCT/US00/22315
; PRIOR FILING DATE: 1999-08-17; 1999-11-09; 2000-08-14
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PERL Program
; SEQ ID NO 1
```

Q.

Db 121 SRALGHENGSLDPEQGMIPENWAPMLAQALE-ALQPALQCLKYKLVFSGRESPEEGEE 179
QY 176 NFDPLWHTNEVLEEWQVSDVEKRRRLMESLRGPAADVIRILKSNPAITTAECCLKALEQ 235
Db 180 EFGRMWFHTTQMIKAWQVDPVEKRRRLMESLRGPAADVIRILKSNPAITTAECCLKALEE 239
QY 236 VFGSVSSRDAQIKFLNTYQNPGEKLSAYVIRLEPLLOKVEKGAKDKNNVQARLEQVI 295
Db 240 VFGVTDPNPRELQVKYLTYYQKDEEKL SAYVIRLEPLLOKLVQVGAERDAVNOARLDQVI 299
QY 296 AGAHSGAIRRQLWLTGAGEGPG 318
Db 300 AGAVHK-TIRRELNPEDGPAPG 321

RESULT 9

US-10-341-434-10
; Sequence 10, Application US/10341434
; Publication No. US20030215835A1
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies
; TITLE OF INVENTION: Differentially Regulated Prostate Cancer Genes
; CURRENT APPLICATION NUMBER: US/10/341,434
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 60/348,164
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: US 60/348,119
; PRIOR FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-341-434-10

Query Match 51.3%; Score 887.5; DB 15; Length 351;
Best Local Similarity 55.4%; Pred. No. 8.1e-81;
Matches 179; Conservative 53; Mismatches 84; Indels 7; Gaps 5;
QY 1 MAMTLEDWCRCMDVNSQRTLLVWGIPVNCDEABIEETLQAAM-POVSYRMLGRMFWRREE 59
Db 1 MTLRLLEDWCRCMDNPNRKALLIAGISQSCSVAIEERLQAGLAPLGEYRLLGEMFRDE 60
QY 60 NAKAALLELTGADVAAIPREMCKGGVWKVLFKPTSDAEFLERHLFLAREGWTVDV 119
Db 61 NRKVALVGLTAETSHALVPKEIPKGGIMRWIFKPPDPDNTFLSRLNEFLAGEGTMVTEL 120
QY 120 ARVLGFQNPNT--PTPG--PMPAEMLNVIILDNVIQPLVESIWKYKRLTLFSGKGHPRAWRG 175
Db 121 SRALGHENGSLDPEQGMIPENWAPMLAQALE-ALQPALQCLKYKLVFSGRESPEEGEE 179
QY 176 NFDPLWHTNEVLEEWQVSDVEKRRRLMESLRGPAADVIRILKSNPAITTAECCLKALEQ 235
Db 180 EFGRMWFHTTQMIKAWQVDPVEKRRRLMESLRGPAADVIRILKSNPAITTAECCLKALEE 239
QY 236 VFGSVSSRDAQIKFLNTYQNPGEKLSAYVIRLEPLLOKVEKGAKDKNNVQARLEQVI 295
Db 240 VFGVTDPNPRELQVKYLTYYQKDEEKL SAYVIRLEPLLOKLVQVGAERDAVNOARLDQVI 299
QY 296 AGAHSGAIRRQLWLTGAGEGPG 318
Db 300 AGAVHK-TIRRELNPEDGPAPG 321

RESULT 10

US-09-804-014A-73
; Sequence 73, Application US/09804014A
; Publication No. US20030064489A1
; GENERAL INFORMATION:
; APPLICANT: Li, Li

; APPLICANT: Padigar, Muralidhara
; APPLICANT: Vernet, Corine
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard
; APPLICANT: Spaderna, Steven
; APPLICANT: Majumder, Kumud
; TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-721 US
; CURRENT APPLICATION NUMBER: US/09/804,014A
; CURRENT FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: 60/188,316
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 60/188,277
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 60/189,139
; PRIOR FILING DATE: 2000-03-14
; PRIOR APPLICATION NUMBER: 60/189,140
; PRIOR FILING DATE: 2000-03-14
; PRIOR APPLICATION NUMBER: 60/190,401
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/190,231
; PRIOR FILING DATE: 2000-03-17
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 73
; LENGTH: 312
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-804-014A-73
Query Match 50.6%; Score 874.5; DB 12; Length 312;
Best Local Similarity 55.9%; Pred. No. 1.4e-79;
Matches 175; Conservative 53; Mismatches 78; Indels 7; Gaps 5;
QY 1 MAMTLEDWCRCMDVNSQRTLLVWGIPVNCDEABIEETLQAAM-POVSYRMLGRMFWRREE 59
Db 1 MTLRLLEDWCRCMDNPNRKALLIAGISQSCSVAIEERLQAGLAPLGEYRLLGEMFRDE 60
QY 60 NAKAALLELTGADVAAIPREMCKGGVWKVLFKPTSDAEFLERHLFLAREGWTVDV 119
Db 61 NRKVALVGLTAETSHALVPKEIPKGGIMRWIFKPPDPDNTFLSRLNEFLAGEGTMVTEL 120
QY 120 ARVLGFQNPNT--PTPG--PMPAEMLNVIILDNVIQPLVESIWKYKRLTLFSGKGHPRAWRG 175
Db 121 SRALGHENGSLDPEQGMIPENWAPMLAQALE-ALQPALQCLKYKLVFSGRESPEEGEE 179
QY 176 NFDPLWHTNEVLEEWQVSDVEKRRRLMESLRGPAADVIRILKSNPAITTAECCLKALEQ 235
Db 180 EFGRMWFHTTQMIKAWQVDPVEKRRRLMESLRGPAADVIRILKSNPAITTAECCLKALEE 239
QY 236 VFGSVSSRDAQIKFLNTYQNPGEKLSAYVIRLEPLLOKVEKGAKDKNNVQARLEQVI 295
Db 240 VFGVTDPNPRELQVKYLTYYQKDEEKL SAYVIRLEPLLOKLVQVGAERDAVNOARLDQVI 299
QY 296 AGAHSGAIRRQL 308
Db 300 AGAVHK-TIRREL 311

RESULT 11

US-09-804-014A-74
; Sequence 74, Application US/09804014A
; Publication No. US20030064489A1
; GENERAL INFORMATION:
; APPLICANT: Li, Li
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Vernet, Corine
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard
; APPLICANT: Spaderna, Steven
; APPLICANT: Majumder, Kumud
; TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-721 US

;; CURRENT APPLICATION NUMBER: US/09/804,014A
;; CURRENT FILING DATE: 2002-04-24
;; PRIOR APPLICATION NUMBER: 60/188,316
;; PRIOR FILING DATE: 2000-03-10
;; PRIOR APPLICATION NUMBER: 60/188,277
;; PRIOR FILING DATE: 2000-03-10
;; PRIOR APPLICATION NUMBER: 60/189,139
;; PRIOR FILING DATE: 2000-03-14
;; PRIOR APPLICATION NUMBER: 60/189,140
;; PRIOR FILING DATE: 2000-03-14
;; PRIOR APPLICATION NUMBER: 60/190,401
;; PRIOR FILING DATE: 2000-03-17
;; PRIOR APPLICATION NUMBER: 60/190,231
;; PRIOR FILING DATE: 2000-03-17
;; NUMBER OF SEQ ID NOS: 75
;; SOFTWARE: Patent in Ver. 2.1
;; SEQ ID NO 74
;; LENGTH: 312
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-804-014A-74

Query Match 50.6%; Score 874.5; DB 12; Length 312;
Best Local Similarity 55.9%; Pred. No. 1.4e-79;
Matches 175; Conservative 53; Mismatches 78; Indels 7; Gaps 5;
QY 1 MAMTLLDWCRCMDVNSQRTLLVWGIPVNCDEAEIETLQAA-POVSYRMIGRMFWREE 59
DB 1 MTLRLLEDWCRCMDNPRKALIAIGISQCSVAETEEALQAGIPLGEYRLIGRMFRDE 60
QY 60 NAKAALLELTGAVDYAAIIPREMPGKGVWVLFKPTSDAEFLERLHLFLAREGWTVDV 119
DB 61 NKKVALVGLTASTHALVPKEIPGKGWVIFKPPDPDNTFLSLNEFLAGEGTVGEL 120
QY 120 ARVLGFQNP--PTFG--PEMPAEMLYILDNVIQPLVESIWKRLTLPFGKHGPRNRG 175
DB 121 SRALCHENGSLDPEQGMIPENWAPMLAQALE-ALQALQCKYKRLRVESGESPEGEE 179
QY 176 NPDWLEHTNEVLEWQVSDVEKRRRLMESLRGPAADVIRILKSNPAITTAECLEKALBO 235
DB 180 EFGRWMEFTTQMIKAWQVDPVDEKRRLESRLGPAADVIRILKSNPAITTAECLEKALBO 239
QY 236 VFGVSSRDQAQIKFLNTYQNPCKLSAYVIRLEPLQKVBKGAIDKDNVNOARLEQVI 295
DB 240 VFGVTDNPRELQVLYTTQKDBEKLISAYVIRLEPLQKVBKGAIDKDNVNOARLDQVI 299
QY 296 AGAHSGAIRRQL 308
DB 300 AGAVHK-TIRREL 311

RESULT 12
US-10-037-860-13
;; Sequence 13, Application US/10037860
;; Publication No. US20020123114A1
;; GENERAL INFORMATION:
;; APPLICANT: Jerome B. Posner
;; APPLICANT: Josep O. Dalmat
;; APPLICANT: Myrna R. Rosenfeld
;; TITLE OF INVENTION: MA FAMILY POLYPEPTIDES AND ANTI-MA
;; TITLE OF INVENTION: ANTIBODIES
;; FILE REFERENCE: 2581.1004-004
;; CURRENT APPLICATION NUMBER: US/10/037,860
;; CURRENT FILING DATE: 2001-01-04
;; PRIOR APPLICATION NUMBER: 09/189,527
;; PRIOR FILING DATE: 1998-11-10
;; NUMBER OF SEQ ID NOS: 14
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 13
;; LENGTH: 463
;; TYPE: PRT
;; ORGANISM: homo sapiens
US-10-037-860-13

Query Match 44.3%; Score 766.5; DB 13; Length 463;
Best Local Similarity 50.2%; Pred. No. 2.2e-68;
Matches 157; Conservative 50; Mismatches 103; Indels 3; Gaps 2;
QY 1 MAMTLLDWCRCMDVNSQRTLLVWGIPVNCDEAEIETLQAA-POVSYRMIGRMFWREE 59
DB 1 MPTLLDWCRCGEHLNTRCMLIIGIPEDCGEDEFETLQACRHGLGRYRIGRMFRREE 60
QY 60 NAKAALLELTGAVDYAAIIPREMPGKGVWVLFKPTSDAEFLERLHLFLAREGWTVDV 119
DB 61 NAKAALLELTGAVDYAAIIPREMPGKGVWVLFKPTSDAEFLERLHLFLAREGWTVDV 120
QY 120 ARVLGFQNP--PTFG--PEMPAEMLYILDNVIQPLVESIWKRLTLPFGKHGPRNRG 177
DB 121 NRVLGSDTNCAPRTVITSPFETWTAQTLGAAVQPLLEQMLYRELKRVFSGNTISIPGALAF 180
QY 178 DPWLEHTNEVLEWQVSDVEKRRRLMESLRGPAADVIRILKSNPAITTAECLEKALBO 237
DB 181 DAWLEHTTEMLQMWQVPEGEKRRRLMESLRGPAADVIRILKSNPAITTAECLEKALBO 240
QY 238 GSVESSRDQAQIKFLNTYQNPCKLSAYVIRLEPLQKVBKGAIDKDNVNOARLEQVI 297
DB 241 GPVESHKTAQVLCCKAYQEAQEKYSSVIRLEPLQKVBKGAIDKDNVNOARLEQVI 300
QY 298 ANHSGAIRRQL 310
DB 301 ATLPDKLRDKL 313

RESULT 13
US-10-408-765A-2385
;; Sequence 2385, Application US/10408765A
;; Publication No. US20040101874A1
;; GENERAL INFORMATION:
;; APPLICANT: Ghosh, Soumitra S.
;; APPLICANT: Fahy, Eoin D.
;; APPLICANT: Zhang, Bing
;; APPLICANT: Gibson, Bradford W.
;; APPLICANT: Taylor, Steven W.
;; APPLICANT: Glenn, Gary M.
;; APPLICANT: Warnock, Dale E.
;; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
;; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
;; FILE REFERENCE: 660088.465
;; CURRENT APPLICATION NUMBER: US/10/408,765A
;; CURRENT FILING DATE: 2003-04-04
;; NUMBER OF SEQ ID NOS: 3077
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 2385
;; LENGTH: 452
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-408-765A-2385

Query Match 43.0%; Score 744; DB 16; Length 452;
Best Local Similarity 46.0%; Pred. No. 3.9e-66;
Matches 143; Conservative 64; Mismatches 100; Indels 4; Gaps 2;
QY 1 MAMTLLDWCRCMDVNSQRTLLVWGIPVNCDEAEIETLQAA-POVSYRMIGRMFWREE 59
DB 5 MALTLLDWCRCMDMDPRKALLIVGIPMECSEVIEQDTWKAGIOPICATRVILGRMERED 64
QY 60 NAKAALLELTGAVDYAAIIPREMPGKGVWVLFKPTSDAEFLERLHLFLAREGWTVDV 119
DB 65 NAKAVIELADTVNYITLPSHIPKGSWEVVKPRNPDPDEFSLRNLFLKDEGRSMTDV 124
QY 120 ARVLGFQNP--PTFG--PEMPAEMLYILDNVIQPLVESIWKRLTLPFGKHGPRNRG 179
DB 125 ARALGC---CSLPAESLDAEVMPQVRSPPLEPPKSNWYKRLKVFSTASPGCEETFD 181
QY 180 WLEHTNEVLEWQVSDVEKRRRLMESLRGPAADVIRILKSNPAITTAECLEKALBO 239

Db 182 WLEQVTEIMPVSEVEKRRRLLESLRGPALSIIMVLQANNDSTVEQCLDALKQIFGD 241

QY 240 VESSRDAQIKFLNTYONPEKLSAVVIRLEPLLOKVVEKGAIDKDNVNOARLEQVIAGAN 299

Db 242 KEDFRASQRFLOTSFKIGERKSTFLLRLEPLLOKAVHKSPUSVNRSTDMIRUKHLLARVA 301

QY 300 HSGAIRRQLWL 310

Db 302 MTPALRGKLEL 312

RESULT 14

US-10-094-749-1978

; Sequence 1978, Application US/10094749

; Publication No. US20030219741A1

; GENERAL INFORMATION:

; APPLICANT: ISOGAI, TAKAO

; APPLICANT: SUGIYAMA, TOMOYASU

; APPLICANT: OTSUKI, TETSUJI

; APPLICANT: WAKAMATSU, AI

; APPLICANT: SATO, HIROYUKI

; APPLICANT: ISHII, SHIZUKO

; APPLICANT: YANAMOTO, JUN-ICHI

; APPLICANT: ISONO, YUUKO

; APPLICANT: HIO, YURI

; APPLICANT: OTSUKA, KAORU

; APPLICANT: NAGAI, KEIICHI

; APPLICANT: IRIE, RYOTARO

; APPLICANT: TAMECHIKA, ICHIRO

; APPLICANT: SEKI, NAOHICO

; APPLICANT: YOSHIKAWA, TSUTOMU

; APPLICANT: OTSUKA, MOTOTYUKI

; APPLICANT: NAGAHARI, KENJI

; APPLICANT: MASUHO, YASUHIKO

; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA

; FILE REFERENCE: 084335/0160

; CURRENT APPLICATION NUMBER: US/10/094,749

; PRIOR FILING DATE: 2002-03-12

; PRIOR APPLICATION NUMBER: 60/350,435

; PRIOR FILING DATE: 2002-01-24

; PRIOR APPLICATION NUMBER: JP 2001-328381

; PRIOR FILING DATE: 2001-09-14

; NUMBER OF SEQ ID NOS: 3381

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 1978

; LENGTH: 399

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-094-749-1978

Query Match 35.9%; Score 620; DB 15; Length 399;

Best Local Similarity 40.9%; Pred. No. 1.2e-53;

Matches 137; Conservative 64; Mismatches 106; Indels 28; Gaps 6;

QY 1 MAMTLEDRCRGMVNSQRTLLVWGPVNCDEAEIETLQAM-POVSYRMLGRMEWREE 59

Db 1 MAVTMLQDWCRWGVNARRGLLIGIPEDCDDAEFQESLEAALRPNGHFTVLGKAFREED 60

QY 60 NAKAALLELTGAVDYAAIPREMKGKGVKVLFPKPTSDAEF--LERLHLFLAREGTVQ 117

Db 61 NATAALVELDRVNYALVPREIFETGTPGNVNVFVPRCSGEEFLGLGRVFFHFEQEGQWVE 120

QY 118 DVARVLGFQNPTEPTPGMEPAEMLNYI--LDNVIQLVESIWKRLTLFSGKGHPRAWRG 175

Db 121 SVAGALG-----VGLRRVCMWLSRGQAVPWVEAVRCQSGUGVFSGRDQPAPGEE 169

QY 176 NFDPLWEHTNEVLEEQ--VSDVEKRRRLMESLRGPADVIRILKSNPNPAITTAECLEKALE 234

Db 170 SFEVLDHTTTEMLHWQGVSERRERRELLSGLRGTAQLQVHALLAENPARTAQDCLAALA 229

QY 235 QVFGSVSESSDAQIKFLNTYONPEKLSAVVIRLEPLLOKVVEKGAIDKDNVNOARLEQV 294

Db 230 QVFGDNESQATIRVKCLTAQQQSGERLSAFVLRLEVLLOKAMEKEALARASADRVLROM 289

QY 295 IAGANHSGAIRRQLWLTCAGEGPGPKPLSVAGADP 329

Db 290 LTRAHLTEPLDEAL-----RKLRMAGRSP 313

RESULT 15

US-10-037-860-11

; Sequence 11, Application US/10037860

; Publication No. US20020123114A1

; GENERAL INFORMATION:

; APPLICANT: Jerome B. Posner

; APPLICANT: Josep O. Dalmau

; APPLICANT: Myrna R. Rosenfeld

; TITLE OF INVENTION: Ma FAMILY POLYPEPTIDES AND ANTI-Ma

; FILE REFERENCE: 2581.1004-004

; CURRENT APPLICATION NUMBER: US/10/037,860

; CURRENT FILING DATE: 2001-01-04

; PRIOR APPLICATION NUMBER: 09/189,527

; PRIOR FILING DATE: 1998-11-10

; NUMBER OF SEQ ID NOS: 14

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 11

; LENGTH: 283

; TYPE: PRT

; ORGANISM: homo sapiens

US-10-037-860-11

Query Match 32.6%; Score 564; DB 13; Length 283;

Best Local Similarity 47.0%; Pred. No. 3.2e-48;

Matches 117; Conservative 47; Mismatches 73; Indels 12; Gaps 4;

QY 83 GKGVKVKLPKPTSDAEFLERLHLFLAREGTVQDVARVLGFQNPTEPTP-----GPMPA 138

Db 3 GKGVKVKVIFKTPNQDTFELERLNLFLKEGQTVSGMFRALGQGVSPATVPFCISPELLA 62

QY 139 EMLNYILDNVIQPLVESIWKRLTLFSGKGHPRAWGNFDPWLHHTNEVLEEWQVSDVEK 198

Db 63 HLLGQAMAHAPQPLL-PMRYRKLRFVSGSAVPAPPEESFEVWLQATEIVKEWPTBAEK 121

QY 199 RRLMESLRGPADVIRILKSNPNPAITTAECLEKALEQVSGSVSSRDAQIKFLNTYQNP 258

Db 122 KRWLAESLRGPALDLMHIVQADNPISVEECLEAFKQVFGSLSRRTAQVRYLKYQEEG 181

QY 259 EKLSAYVIRLEPLLOKVVEKGAIDKDNVNOARLEQVIAGANHSGAIRRQLWL---TGAGE 315

Db 182 EKVSAYVLRLETLRLKAVEKRAIPRRITADQVRLEQVWAGA---TLNQMLWCLRLRELKDQ 237

QY 316 GPGPKPLSV 324

Db 238 GPPPSFLEL 246

Search completed: September 21, 2004, 14:04:06

Job time : 90.4672 secs

Blank

RESULT 2

US-09-189-527-13
; Sequence 13, Application US/09189527A
; Patent No. 6387639
; GENERAL INFORMATION:
; APPLICANT: Jerome B. Posner
; APPLICANT: Josep O. Dalmau
; APPLICANT: Myrna R. Rosenfeld
; TITLE OF INVENTION: Ma Family Polypeptides and Anti-Ma
; FILE REFERENCE: SLK98-01
; CURRENT APPLICATION NUMBER: US/09/189,527A
; CURRENT FILING DATE: 1998-11-10
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 13
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-189-527-13

Query Match 43.2%; Score 747.5; DB 4; Length 462;
Best Local Similarity 49.8%; Pred. No. 1.1e-73;
Matches 153; Conservative 49; Mismatches 102; Indels 3; Gaps 2;

QY 7 EDWCRGMDVNSORTLLVWGIPVNCDEAEIETLQAAAPQVS-YRMLGRMFWRRENAKAL 65
DB 1 QDWCRGMDVNSORTLLVWGIPVNCDEAEIETLQAAAPQVS-YRMLGRMFWRRENAKAL 60
QY 66 LEITGADVAAIPREMPGKGGVWVLFKPPPTSDAEFLERLHLFLAREGWTQDVAVLGF 125
DB 61 LELAQIDVALLPREIPGKGPWEVIVKPNRSDGEFLNRLNRELEERTVSDMNVLS 120
QY 126 QNPTPTPGEMPAMLYN--ILDNVIQPLVESIWKRLTLFSGKGHPRAWGNFDPWLEH 183
DB 121 DTNCSAPRVTSPEFTWAGTLLAAVQPLLEQMLYRELRFVSGNTISIFCALAFDAWLEH 180
QY 184 TNEVLEEMQVDEKERRLMESLRGAADVIRILKSNPAITTAECLEKALEQVFGSVES 243
DB 181 TTEMLQMWQVPEGEKRRRLMECLRGALQVSGLRASNASITVECLALQVFGPVESH 240
QY 244 RDAQIKFLNTYQNGEKLKSAVIRLEPLQKVEKGAIDKDNVNOARLEQVIAGANHSGA 303
DB 241 KIAQVKLCRAYQAGEKVSFVLRLEPLQRAVENNVSRNNVQTLRLKVLSGATLPDK 300
QY 304 IRRQLWL 310
DB 301 LRDKLKL 307

RESULT 3

US-09-189-527-7
; Sequence 7, Application US/09189527A
; Patent No. 6387639
; GENERAL INFORMATION:
; APPLICANT: Jerome B. Posner
; APPLICANT: Josep O. Dalmau
; APPLICANT: Myrna R. Rosenfeld
; TITLE OF INVENTION: Ma Family Polypeptides and Anti-Ma
; FILE REFERENCE: SLK98-01
; CURRENT APPLICATION NUMBER: US/09/189,527A
; CURRENT FILING DATE: 1998-11-10
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 195
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-189-527-7

Query Match 25.5%; Score 441; DB 4; Length 195;
Best Local Similarity 46.4%; Pred. No. 2.1e-40;

Matches 90; Conservative 37; Mismatches 61; Indels 6; Gaps 4;
QY 3 MTLLEDWCRGMDVNSORTLLVWGIPVNCDEAEIETLQAAAPQVS-YRMLGRMFWRRENA 61
DB 2 LALLEDWCRIMSDQKSLMVTGIPADFEAEIQEVLQETLSLGRYLLGLKFRQENA 61
QY 62 KAALELTGADVAAIPREMPGKGGVWVLFKPPPTSDAEFLERLHLFLAREGWTQDVAV 121
DB 62 NAVLELLEDTDVSAIPSEVQKGGVWVLFKPPPTSDAEFLERLHLFLAREGWTQDVAV 121
QY 122 VLGFQ--NPTPTPG--GPEMPAEMLYILDNVIQPLVESIWKRLTLFSGKGHPRAWGNF 177
DB 122 ALGQEALSPATVPCISPELLAHLQAAHAPOPLL-PMRYRKLRFVSGSAVPAPEERSF 180
QY 178 DPMLEHTNEVLEEM 191
DB 181 EVMLEQATEIVKRW 194

RESULT 4

US-09-914-259-25
; Sequence 25, Application US/09914259
; Patent No. 6495336
; GENERAL INFORMATION:
; APPLICANT: Makowski, Lee
; APPLICANT: Hyman, Paul
; APPLICANT: Williams, Mark
; TITLE OF INVENTION: STAGED ASSEMBLY OF NANOSTRUCTURES
; FILE REFERENCE: 8471-010-999
; CURRENT APPLICATION NUMBER: US/09/914,259
; CURRENT FILING DATE: 2000-11-21
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 935
; TYPE: PRT
; ORGANISM: Syncephalastrum racemosum
US-09-914-259-25

Query Match 6.0%; Score 103; DB 4; Length 935;
Best Local Similarity 21.7%; Pred. No. 0.065;

Matches 69; Conservative 56; Mismatches 121; Indels 72; Gaps 15;
QY 16 NSORTLLVWGIPVNCDEAEIETLQAAAPQVS-YRMLGRMFWRRENAKAL 68
DB 297 NSRTTILINCSPPSSYNEAETLSTLRFGARAKSIKNKAKV-----NADLSPAELKALIKV 351
QY 69 -TGADVAAIPREMPGKGGVWVLFKPPPTSDAEFLERLHLFLAREG-WTVQDVAVLGFQ 126
DB 352 KSEAVTYQTYIAALEGEVNVWRTGGTVP-----EGKWVTMDKVGSGDFA 395
QY 127 NPTPTPGPEMPAEMLYILDNVIQPLVESIWKRLTLFSGKGHPRAWGNFDPWLEHTNE 186
DB 396 GLPPAPGFKSP-----VSDEGSRPATPV-----PTLEKDEREERFIKRENE 435
QY 187 VLEEWQVSDVE-----KRRIMESLR--GPAADVIRILKSNPAIT--AECLKALEQVF 237
DB 436 LMD--QISEKETELTNREKLLESIREMGVYKEQSVTKENQOQMTSELRLQLQKV-- 492
QY 238 GSVESSRDAQIKFLNTYQNGEKLKSAVIRLEPLQKVEKGAIDKDNVNOARLEQ 293
DB 493 -SYESKENAIT--VDSLKEANQDLMALEELKKNLSEMRQAHKDATDSDKEKRAEKMAQ 549
QY 294 VIAGANHSGAI---RRQL 308
DB 550 MMSGFDPGSLINDKERQI 567

RESULT 5

US-09-368-590-2
; Sequence 2, Application US/09368590
; Patent No. 6187563
; GENERAL INFORMATION:

;; CITY: Falls Church
;; STATE: Virginia
;; COUNTRY: USA
;; ZIP: 22040-0747
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent in Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/07/920,281C
;; FILING DATE: 13-AUG-1992
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Murphy Jr., Gerald M.
;; REGISTRATION NUMBER: 28,977
;; REFERENCE/DOCKET NUMBER: 828-103P
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 703-241-1300
;; TELEFAX: 703-241-2848
;; TELEX: 248345
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 2431 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-07-920-281C-2

Query Match 5.3%; Score 92.5; DB 1; Length 2431;
Best Local Similarity 19.6%; Pred. No. 4.6;
Matches 74; Conservative 50; Mismatches 112; Indels 141; Gaps 15;

QY 52 GRMFWREENAKAALBELTGAVDYAALPREMPGGGKVKVLFKPPPTSDAEFLERLHLFL-- 109
Db 865 GKMRITNPNCKPIIIDTTG-----QTKPKPGDIVLTCFRGWAKQLQLDYRGHEVMTA 916

QY 110 -AREGWTVDQVAVLGFQNPPTPGPEMPAEMLYINDVNIQPLVESIWYKRLTLFSGKG 168
Db 917 AASQGLTRGVAVRQKVNENPLYAP--ASEHVNLLTRTEDRLV-----WKTILA----- 964

QY 169 HPRAMRGNDPWLHETNEV-----LEEWQVSDVEKRRRLMESLRGPAADV----- 213
Db 965 -----GDPWKVLSNIPQGNFTATLEWQ-----EEHDKIMKVIEGPAAPVDAFONKA 1012

QY 214 -----IRILKSNNPATTA-----ECLKALEQV----- 236
Db 1013 NVCWAKSLVPDLDTAGIRLTAEWSTIITAFKEDRAYSPVVALNEICTKYGYVDLDSGLF 1072

QY 237 -----FGSVESRR-DAQIKFLNTYONPGEKLSAVVIRLEPL 271
Db 1073 SAPKVSLLYENNHNDRPGRMVGFNAATAARLEAHTFLKGQWHTGKQAVIAERKIQPL 1132

QY 272 --LQKVE-----KGAIDKDNVNOARLEQVIAGANHSgai--RRQLWLT 311
Db 1133 SVLDNVIPINRRLPHALVAEYKTVKGRVWLVNKGVRGYHLLVSEYNLALPRRRVTWLS 1192

QY 312 GAGEGPGPKPLSVAGAD 328
Db 1193 -----PLNVTGAD 1200

RESULT 9
US-08-466-277-2
; Sequence 2, Application US/08466277
; Patent No. 6190666
; GENERAL INFORMATION:
; APPLICANT: Garoff, Henrik
; Liljestrom, Peter
; TITLE OF INVENTION: DNA Expression Systems Based on
; Alphaviruses
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:

;; ADDRESSEE: Birch, Stewart, Kolasch & Birch
;; STREET: P.O. Box 747
;; CITY: Falls Church
;; STATE: Virginia
;; COUNTRY: USA
;; ZIP: 22040-0747
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent in Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/466,277
;; FILING DATE: 06-Jun-1995
;; CLASSIFICATION: <Unknown>
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 07/920,281
;; FILING DATE: <Unknown>
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Murphy Jr., Gerald M.
;; REGISTRATION NUMBER: 28,977
;; REFERENCE/DOCKET NUMBER: 828-103P
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 703-241-1300
;; TELEFAX: 703-241-2848
;; TELEX: 248345
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 2431 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-08-466-277-2

Query Match 5.3%; Score 92.5; DB 3; Length 2431;
Best Local Similarity 19.6%; Pred. No. 4.6;
Matches 74; Conservative 50; Mismatches 112; Indels 141; Gaps 15;

QY 52 GRMFWREENAKAALBELTGAVDYAALPREMPGGGKVKVLFKPPPTSDAEFLERLHLFL-- 109
Db 865 GKMRITNPNCKPIIIDTTG-----QTKPKPGDIVLTCFRGWAKQLQLDYRGHEVMTA 916

QY 110 -AREGWTVDQVAVLGFQNPPTPGPEMPAEMLYINDVNIQPLVESIWYKRLTLFSGKG 168
Db 917 AASQGLTRGVAVRQKVNENPLYAP--ASEHVNLLTRTEDRLV-----WKTILA----- 964

QY 169 HPRAMRGNDPWLHETNEV-----LEEWQVSDVEKRRRLMESLRGPAADV----- 213
Db 965 -----GDPWKVLSNIPQGNFTATLEWQ-----EEHDKIMKVIEGPAAPVDAFONKA 1012

QY 214 -----IRILKSNNPATTA-----ECLKALEQV----- 236
Db 1013 NVCWAKSLVPDLDTAGIRLTAEWSTIITAFKEDRAYSPVVALNEICTKYGYVDLDSGLF 1072

QY 237 -----FGSVESRR-DAQIKFLNTYONPGEKLSAVVIRLEPL 271
Db 1073 SAPKVSLLYENNHNDRPGRMVGFNAATAARLEAHTFLKGQWHTGKQAVIAERKIQPL 1132

QY 272 --LQKVE-----KGAIDKDNVNOARLEQVIAGANHSgai--RRQLWLT 311
Db 1133 SVLDNVIPINRRLPHALVAEYKTVKGRVWLVNKGVRGYHLLVSEYNLALPRRRVTWLS 1192

QY 312 GAGEGPGPKPLSVAGAD 328
Db 1193 -----PLNVTGAD 1200

RESULT 10
US-09-055-097-1
; Sequence 1, Application US/09055097
; Patent No. 5955282
; GENERAL INFORMATION:

; APPLICANT: Hillman, Jennifer L.
 ; APPLICANT: Guegler, Karl J.
 ; APPLICANT: Corley, Neil C.
 ; APPLICANT: Shah, Purvi
 ; APPLICANT: Patterson, Chandra
 ; TITLE OF INVENTION: HUMAN OXIDIZED LDL RECEPTOR
 ; NUMBER OF SEQUENCES: 7
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Incyte Pharmaceuticals, Inc.
 ; STREET: 3174 Porter Drive
 ; CITY: Palo Alto
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94304
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; FILING DATE: Filed Herewith
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Cerrone, Michael C.
 ; REGISTRATION NUMBER: 39,132
 ; REFERENCE/DOCKET NUMBER: PF-0490 US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (650) 855-0555
 ; TELEFAX: (650) 845-4166
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 378 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; IMMEDIATE SOURCE:
 ; LIBRARY: DUODNOT02
 ; CLONE: 1705085
 ; US-09-055-097-1

Query Match 5.3%; Score 91.5; DB 2; Length 378;
 Best Local Similarity 23.9%; Pred. No. 0.27;
 Matches 61; Conservative 26; Mismatches 89; Indels 79; Gaps 15;
 Qy 11 RGMV-----NSQRTLLVW---GIPVNCDEAEIETLQAAMPQVSYSR-----MLGR 53
 Db 101 RGLRVQTLFLGEPNACHP--VWGSQSDLASESAQGDILQAAP-QDSYRNLTLLKTLGG 157
 Qy 54 MFWEEENAKAALLELTGAVD-YAALPREMP-----GKGWVKVLFKPP-TSDAEFLERL 105
 Db 158 LNWAEKHCPMARYLVKTDVVVNVVNPVSELVLRGGWGWQWERTPEQREAEQGGQVL 217
 Qy 106 H-----LFLAREGWTQDVVARVLGFQNPPTPGPEMPAEMLYLDNVIQPLVESIWK 159
 Db 218 HSEVPLLYLGRVHVRV-----NPSRTPGGR-----HRVSEEQW-- 251
 Qy 160 RLTLFSGKHPRAMGNFDPMLHETNEVLEWQVS---DVEKRRRLMESLRGPAADVIRI 216
 Db 252 -----PHTW-GPPPPYASGTGYLSASAVQLILKVASRAPLL-----PLEDVFGV 295
 Qy 217 LKSNNPALTITAECLK 231
 Db 296 VSARRGGLAPTQCCKV 310

RESULT 11

US-09-373-902-1

; Sequence 1, Application US/09373902

; Patent No. 6649737

; GENERAL INFORMATION:
 ; APPLICANT: Hillman, Jennifer L.
 ; Guegler, Karl J.
 ; Corley, Neil C.
 ; Shah, Purvi
 ; Patterson, Chandra
 ; TITLE OF INVENTION: HUMAN OXIDIZED LDL RECEPTOR
 ; NUMBER OF SEQUENCES: 7
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Incyte Pharmaceuticals, Inc.
 ; STREET: 3174 Porter Drive
 ; CITY: Palo Alto
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94304
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/373,902
 ; FILING DATE: 12-Aug-1999
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/055,097
 ; FILING DATE: <Unknown>
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Cerrone, Michael C.
 ; REGISTRATION NUMBER: 39,132
 ; REFERENCE/DOCKET NUMBER: PF-0490 US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (650) 855-0555
 ; TELEFAX: (650) 845-4166
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 378 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; IMMEDIATE SOURCE:
 ; LIBRARY: DUODNOT02
 ; CLONE: 1705085
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
 ; US-09-373-902-1

Query Match 5.3%; Score 91.5; DB 4; Length 378;
 Best Local Similarity 23.9%; Pred. No. 0.27;
 Matches 61; Conservative 26; Mismatches 89; Indels 79; Gaps 15;
 Qy 11 RGMV-----NSQRTLLVW---GIPVNCDEAEIETLQAAMPQVSYSR-----MLGR 53
 Db 101 RGLRVQTLFLGEPNACHP--VWGSQSDLASESAQGDILQAAP-QDSYRNLTLLKTLGG 157
 Qy 54 MFWEEENAKAALLELTGAVD-YAALPREMP-----GKGWVKVLFKPP-TSDAEFLERL 105
 Db 158 LNWAEKHCPMARYLVKTDVVVNVVNPVSELVLRGGWGWQWERTPEQREAEQGGQVL 217
 Qy 106 H-----LFLAREGWTQDVVARVLGFQNPPTPGPEMPAEMLYLDNVIQPLVESIWK 159
 Db 218 HSEVPLLYLGRVHVRV-----NPSRTPGGR-----HRVSEEQW-- 251
 Qy 160 RLTLFSGKHPRAMGNFDPMLHETNEVLEWQVS---DVEKRRRLMESLRGPAADVIRI 216
 Db 252 -----PHTW-GPPPPYASGTGYLSASAVQLILKVASRAPLL-----PLEDVFGV 295
 Qy 217 LKSNNPALTITAECLK 231
 Db 296 VSARRGGLAPTQCCKV 310

RESULT 12

US-09-543-681A-5919

; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 11551
; LENGTH: 178
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-11551

Query Match	5.1%;	Score 89;	DB 4;	Length 178;
Best Local Similarity	22.2%;	Pred. No. 0.15;		
Matches	45;	Conservative 31;	Mismatches 71;	Indels 56; Gaps 7;

Qy	23	VMGIPVNCDEAEIETLQAMPQVSRYMLGRMFWEENAKAALLELTGAVDYAAIPREMP	82
Db	13	LFGLPAS-----YTLSEQLAVRYQDLQRYHPDKFASAPAEQLAAVQHSATINQ--	63
Qy	83	GKGGVWKVLFKPPTSDAEFLERLHLF-LAREGWTVDVARVLGFQNPFTTGPPEMPAEML	141
Db	64	----AWQTL-RHPLTRAEYLLSLHGFDLASEQHTVRDTAFL-----MEQLEL	105
Qy	142	NVILDNVIQ-----PLVESIWYKRLTLFSGKGHPRAWRGNFPDPWLEHTNEVLEE	190
Db	106	REELDEICQAKDDARLEGFIKRVKALFDTROQLMVDQLHNSWEA-----	150
Qy	191	WQVSDVEXRRRLMESLRGPAADV	213
Db	151	--AADTVKRLRFLDKLRSSAEEL	171

Search completed: September 21, 2004, 13:36:16
Job time : 26.461 secs

Blank

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 13:20:21 : Search time 60.7447 Seconds
(without alignments)
1530.308 Million cell updates/sec

Title: US-10-037-860-4

Perfect score: 1729

Sequence: 1 MAMTLLDWCGRMDVNSQRT.....LTGAGEGFGPKELSVAGADP 329

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*

- 1: geneseqp1980s:*
- 2: geneseqp1990s:*
- 3: geneseqp2000s:*
- 4: geneseqp2001s:*
- 5: geneseqp2002s:*
- 6: geneseqp2003as:*
- 7: geneseqp2003bs:*
- 8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1704	98.6	329	3 AAB12525	Aab12525 Human Ma1
2	1602	92.7	353	4 AAB74701	Aab74701 Human mem
3	887.5	51.3	351	4 AAE01340	Aae01340 Human gen
4	887.5	51.3	351	4 AAB74695	Aab74695 Human mem
5	887.5	51.3	351	4 AAU08664	Aau08664 Human ORF
6	874.5	50.6	312	3 AAB43023	Aab43023 Human ORF
7	770.5	44.6	364	7 ADC08977	Adc08977 Oncogene
8	766.5	44.3	455	5 ABB05727	Abb05727 Human sig
9	766.5	44.3	463	3 AAB12529	Aab12529 Human Ma5
10	766.5	44.3	463	3 AAB42315	Aab42315 Human ORF
11	744	43.0	452	6 AAO16179	Aao16179 Human pro
12	620	35.9	399	6 ADA54410	Ada54410 Human pro
13	620	35.9	399	6 ABG99947	Abg99947 Human nov
14	570.5	33.0	280	4 AAE01336	Aae01336 Human gen
15	565	32.7	283	3 AAB12528	Aab12528 Human Ma4
16	540	31.2	237	4 AAB94854	Aab94854 Human Ma4
17	441	25.5	195	3 AAB12526	Aab12526 Human Ma2
18	399	23.1	110	4 AAO01787	Aao01787 Human pol
19	391.5	22.6	439	6 ABP97536	Abp97536 Human sec
20	335.5	19.4	403	5 ABG97495	Abg97495 Human NOV
21	335.5	19.4	403	5 AAM51624	Aam51624 KIAA0883-
22	335.5	19.4	403	6 ABO14772	Abo14772 Novel hum
23	335	19.4	403	6 ABO14773	Abo14773 Novel hum
24	332	19.2	402	4 AAB60478	Aab60478 Human cel
25	326	18.9	337	4 AAM25693	Aam25693 Human pro

ALIGNMENTS

RESULT 1

AAB12525
ID AAB12525 standard; protein; 329 AA.

XX AAB12525;

DT 02-NOV-2000 (first entry)

XX Human Ma1 protein SEQ ID NO:4.

XX Ma1; Ma2; Ma3; Ma4; Ma5; paraneoplastic limbic encephalitis; diagnosis;
KW paraneoplastic syndrome; neoplasm; colon cancer; brainstem encephalitis;
KW breast cancer; parotid gland cancer; lung cancer; testicular cancer;
KW germ-cell tumour.

XX Homo sapiens.

XX JP2000146982-A.

XX 26-MAY-2000.

XX 10-NOV-1999; 99JP-00320171.

XX 10-NOV-1998; 98US-00189527.

XX (SLOK) SLOAN KETTERING INST CANCER RES.

XX WPI; 2000-468119/41.

XX N-PSDB; AAA60833.

PT Diagnosing paraneoplastic syndrome e.g. paraneoplastic limbic
PT encephalitis or neoplasm e.g. colon cancer comprising assessing a test
PT sample for the presence or absence of antibodies to a Ma family
polypeptide.

XX Claim 48; Fig 1; 27pp; Japanese.

CC The present invention describes a method for diagnosing a paraneoplastic
CC syndrome or neoplasm. The method comprises assessing a test sample for
CC the presence or absence of antibodies to a Ma family polypeptide (I). The
CC method is used to diagnose a paraneoplastic syndrome especially
CC paraneoplastic limbic encephalitis and/or brainstem encephalitis or
CC neoplasm in an individual. The method diagnoses the neoplasm by assessing
CC antibodies to (I) preferably Ma1, which is indicative presence of breast
CC cancer, colon cancer, parotid gland cancer, lung cancer, testicular
CC cancer and germ-cell tumours or Ma2, which is indicative of testicular
CC cancer, germ-cell tumour, and lung cancer. The present sequence is the
CC Ma1 protein as given in the present invention

26	256.5	14.8	149	3	AAB12527	Aab12527 Human Ma3
27	246.5	14.3	116	4	AAM14937	Aam14937 Peptide #
28	246.5	14.3	116	4	AB333906	Ab333906 Peptide #
29	246.5	14.3	116	4	AAM27367	Aam27367 Peptide #
30	246.5	14.3	116	4	ABB28727	Abb28727 Peptide #
31	246.5	14.3	116	4	ABB19347	Abb19347 Protein #
32	246.5	14.3	116	4	AAM67075	Aam67075 Human bon
33	246.5	14.3	116	4	AAM54673	Aam54673 Human bra
34	246.5	14.3	116	4	ABG48741	Abg48741 Human liv
35	246.5	14.3	116	4	AAM02665	Aam02665 Peptide #
36	246.5	14.3	116	5	ABG36734	Abg36734 Human pep
37	221	12.8	615	4	ABG19651	Abg19651 Novel hum
38	221	12.8	615	4	ABG14259	Abg14259 Novel hum
39	192	11.1	538	4	ABU52641	Abu52641 Human bra
40	192	11.1	538	7	ADC31124	Adc31124 Human nov
41	133	7.7	584	4	AAU28186	Aau28186 Novel hum
42	115.5	6.7	555	6	ABU21717	Abu21717 Protein e
43	114.5	6.6	555	6	ABU22846	Abu22846 Protein e
44	114	6.6	373	3	AAV75199	Aav75199 Neisseria
45	114	6.6	558	6	ABU37227	Abu37227 Protein e

```
XX SQ Sequence 329 AA;
Query Match 98.6%; Score 1704; DB 3; Length 329;
Best Local Similarity 98.8%; Pred. No. 1.4e-165;
Matches 325; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 MAMTLLDWCRCGMDVNSQRTLLVWGIPVNCDEAEIETTLQAAMPQVSYRMLGRMPWREN 60
DB 1 MAMTLLDWCRCGMDVNSQRTLLVWGIPVNCDEAEIETTLQAAMPQVSYRMLGRMPWREN 60

QY 61 AKAAELLELTGAVDYAAIPREMPCKGGVWKVLPKPTSDAEFLERLHLFLAREGWTVDVA 120
DB 61 AKAAELLELTGAVDYAAIPREMPCKGGVWKVLPKPTSDAEFLERLHLFLAREGWTVDVA 120

QY 121 RVLGFGONPTPTGPEMPAEMLYILDNVIQPLVESIWKYKRLTLFSKGHPRAWGNFDPW 180
DB 121 RVLGFGONPTPTGPEMPAEMLYILDNVIQPLVESIWKYKRLTLFSKGHPRAWGNFDPW 180

QY 121 RVLGFGONPTPTGPEMPAEMLYILDNVIQPLVESIWKYKRLTLFSKGHPRAWGNFDPW 180
DB 121 RVLGFGONPTPTGPEMPAEMLYILDNVIQPLVESIWKYKRLTLFSKGHPRAWGNFDPW 180

QY 181 LEHTNEVLEWQVSDVEKRRRLMESLRGPAADVIRILKSNPAITTAECLEQVFGSV 240
DB 181 LEHTNEVLEWQVSDVEKRRRLMESLRGPAADVIRILKSNPAITTAECLEQVFGSV 240

QY 241 ESSRDAQIKFLNTYQNPGEKLSAYVIRLEPLQKVVEKGAIDKDNVNQARLEQVIAGANH 300
DB 241 ESSRDAQIKFLNTYQNPGEKLSAYVIRLEPLQKVVEKGAIDKDNVNQARLEQVIAGANH 300

QY 301 SGAIRQLWLTGAGEGPGPKPLSVAGADP 329
DB 301 SGAIRQLWLTGAGEGPGPKPLSVAGADP 329

RESULT 2
AAB74701
ID AAB74701 standard; protein; 353 AA.
AC AAB74701;
XX
XX 12-JUN-2001 (first entry)
XX
XX Human membrane associated protein MEMAP-7.
DE
XX Human; membrane associated protein; MEMAP; diagnosis; cytostatic;
KW antiinflammatory; anticonvulsant; immunosuppressive; antidiarrheic;
KW antiarteriosclerotic; gene therapy; cell proliferative disorder;
KW autoimmune disorder; inflammatory disorder; neurological disorder;
KW gastrointestinal disorder; cancer; inflammation; atherosclerosis;
KW epilepsy; diarrhoea.
XX
XX Homo sapiens.
OS
XX WO200112662-A2.
PN
XX 22-FEB-2001.
PD
XX
XX 14-AUG-2000; 2000WO-US022315.
PF
XX
XX 17-AUG-1999; 99US-0149641P.
PR
XX 09-NOV-1999; 99US-0164203P.
XX
XX (INCY-) INCYTE GENOMICS INC.
PA
XX
XX Lal P, Yue H, Tang YT, Bandman O, Burford N, Azimzai Y;
PI Baughn MR, Lu DAM, Patterson C;
XX
XX WPI; 2001-168860/17.
DR N-PSDB; AAF81747.
XX
XX Isolated polypeptide with a human membrane associated protein sequence is
PT useful for the diagnosis, prevention and treatment of cell proliferative,
PT autoimmune/inflammatory, neurological and gastrointestinal disorders.
XX
XX Claim 1; Page 119-120; 173pp; English.
PS
```

```
XX SQ Sequence 353 AA;
Query Match 92.7%; Score 1602; DB 4; Length 353;
Best Local Similarity 96.6%; Pred. No. 4.4e-155;
Matches 308; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

QY 1 MAMTLLDWCRCGMDVNSQRTLLVWGIPVNCDEAEIETTLQAAMPQVSYRMLGRMPWREN 60
DB 1 MAMTLLDWCRCGMDVNSQRTLLVWGIPVNCDEAEIETTLQAAMPQVSYRMLGRMPWREN 60

QY 61 AKAAELLELTGAVDYAAIPREMPCKGGVWKVLPKPTSDAEFLERLHLFLAREGWTVDVA 120
DB 61 AKAAELLELTGAVDYAAIPREMPCKGGVWKVLPKPTSDAEFLERLHLFLAREGWTVDVA 120

QY 121 RVLGFGONPTPTGPEMPAEMLYILDNVIQPLVESIWKYKRLTLFSKGHPRAWGNFDPW 180
DB 121 RVLGFGONPTPTGPEMPAEMLYILDNVIQPLVESIWKYKRLTLFSKGHPRAWGNFDPW 180

QY 181 LEHTNEVLEWQVSDVEKRRRLMESLRGPAADVIRILKSNPAITTAECLEQVFGSV 240
DB 181 LEHTNEVLEWQVSDVEKRRRLMESLRGPAADVIRILKSNPAITTAECLEQVFGSV 240

QY 241 ESSRDAQIKFLNTYQNPGEKLSAYVIRLEPLQKVVEKGAIDKDNVNQARLEQVIAGANH 300
DB 241 ESSRDAQIKFLNTYQNPGEKLSAYVIRLEPLQKVVEKGAIDKDNVNQARLEQVIAGANH 300

QY 301 SGAIRQLWLTGAGEGPGP 319
DB 301 SGAIRQLWLTGAGEGPGP 319

RESULT 3
AAE01340
ID AAE01340 standard; protein; 351 AA.
XX
XX AAE01340;
XX
XX 17-JUL-2001 (first entry)
XX
XX Human gene 22 encoded secreted protein fragment, SEQ ID NO:205.
XX
XX Human; secreted protein; proliferative disorder; cancer; tumour;
KW foetal abnormality; developmental abnormality; haematopoietic disorder;
KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
KW inflammation; allergy; neurological disorder; Alzheimer's disease;
KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;
KW skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
KW cardiovascular disorder; angiogenic disorder; kidney disorder;
KW gastrointestinal disorder; pregnancy-related disorder;
XX
```


CC prevent a disorder associated with MEMAP. Polynucleotides encoding MEMAP
CC can be delivered to target cells with genetic abnormalities with respect
CC to the expression of MEMAP to treat or prevent a disorder associated with
CC MEMAP

XX Sequence 351 AA;
Query Match 51.3%; Score 887.5; DB 4; Length 351;
Best Local Similarity 55.4%; Pred. No. 7.7e-82;
Matches 179; Conservative 53; Mismatches 84; Indels 7; Gaps 5;
QY 1 MAMTLLDWCRCGMDVNSQRTLLVWGPVNCDEAEIBETLQAAM-PQVSVMGLGRMFWRREE 59
Db 1 MTLRLLEDWCRCGMDVNSQRTLLVWGPVNCDEAEIBETLQAAM-PQVSVMGLGRMFWRREE 60
QY 60 NAKAALLEITGAVDYAAIPREMPKGGVWVLPKPTSDAEFLERLHLFLAREGWTVDV 119
Db 61 NRKVALVGLTAETSHALVPKEIPKGGIWRVIFKPPDPNTFLSRLNEFLAGEMTVGEL 120
QY 120 ARVLGFQNPPT--PTPG--PEMPAEMLNLYILDVNIQPLVESIWKLTFLSGKHPRAWRG 175
Db 121 SRAIGHENGSLDPEQCMIPEMWAPMLAQALE-ALQPALQCLKYKLVRSGRSPFEGEE 179
QY 176 NFDPMLEHTNEVLEEQVSDVEKRRRLMESLRGPAADVIRILKNNPAITTAECLEKALEQ 235
Db 180 EFGRWPFHTTQMIKAMQVDPVDEKRRRLLESRLRGPAADVIRILKNNPLITVDCECLQALEE 239
QY 236 VFGSVSSRDQAKFLNTYQNPCKLSAVIRLEPLLOKVEKGAIDKONVQARLEQVI 295
Db 240 VFGVTDNPRELQVKYLTYYQKDEKLSAVYLRLEPLLOKLVORGAIERDAVNOARLDQVI 299
QY 296 AGANHSGAIRRQLWLTGAGEGPG 318
Db 300 AGAVHK-TIRRELNPEDGPAFG 321

RESULT 5
AAU08664
ID AAU08664 standard; protein; 351 AA.
XX AAU08664;
AC AAU08664;
XX 18-DEC-2001 (first entry)
XX Human NOV8 protein.
XX Human; NOV8; cytostatic; neutropenic; neuroprotective; vulnary;
KW cerebroprotective; antiparkinsonian; hypotensive; antidiabetic;
KW antidiabetic; antipsoriatic; antiinflammatory; immunosuppressive;
KW antithrombotic; dermatological; cancer; neurological disorder;
KW Parkinson's disease; diabetes mellitus; asthma; enamel defect;
KW immune disorder; autoimmune disease; respiratory disorder; bone disorder;
KW musculoskeletal disorders; leukaemia; lymphoma;
KW cell growth regulation disorder; lesional psoriatic skin;
KW atherosclerosis; abdominal aortic aneurysm.
XX Homo sapiens.
XX WC200168851-A2.
XX 20-SEP-2001.
XX 12-MAR-2001; 2001WO-US007735.
XX 10-MAR-2000; 2000US-0188277P.
XX 10-MAR-2000; 2000US-0188316P.
XX 14-MAR-2000; 2000US-0189139P.
XX 14-MAR-2000; 2000US-0189140P.
XX 17-MAR-2000; 2000US-0190231P.
XX 17-MAR-2000; 2000US-0190401P.
XX (CURA-) CURAGEN CORP.
XX

PI Padigaru M, Vernet CAM, Fernandes E, Shimkets RA, Spaderina SK;
PI Majumder K, Li L;
XX WPI: 2001-570869/64.
DR N-PSDB; AAS13342.
XX Novel polypeptides and nucleic acids homologous to members of collagen,
PT potassium channel, tuftelin family of proteins for diagnosing, treating
PT cancer, atherosclerosis, neurological, skin and enamel defect disorders.
XX Claim 1; Page 29; 128pp; English.
XX The invention relates to isolated NOVX (NOVX1-11) polypeptides and the
CC polynucleotides that encode them. NOVX polypeptides, polynucleotides and
CC anti-NOVX antibodies are useful for treating or preventing a pathology
CC associated with NOVX polypeptide in humans and for treating a syndrome
CC associated with human disease e.g. disorders characterised by altered
CC cell motility, proliferation and migration e.g. cancer, angiogenesis and
CC wound healing (NOV1-3), neurological disorders, e.g. episodic ataxia,
CC autosomal dominant myokymia, stroke, Parkinson's disease, Alzheimer's
CC disease, non-insulin dependent diabetes mellitus, asthma, hypertension
CC and seizure (NOV4), enamel defects, such as amelogenesis imperfecta and
CC disorders involving enamel defects, including hypoplasia and
CC hypomineralisation (NOV5-7), paraneoplastic neurological disorders, e.g.
CC paraneoplastic limbic of brain-stem encephalitis occurring during
CC testicular cancer, diabetes, reproductive health, metabolic and endocrine
CC disorders, gastrointestinal disorders, immune disorders and autoimmune
CC diseases, respiratory disorders, bone disorders, musculoskeletal
CC disorders, leukaemia/lymphoma and tissue/cell growth regulation disorders
CC (NOV8), lesional psoriatic skin (NOV9-10) and atherosclerosis, abdominal
CC aortic aneurysm and neurological disorders (NOV11). NOVX polypeptide is
CC also useful for identifying an agent that binds to it and a cell
CC expressing NOVX polypeptide is useful for identifying a therapeutic agent
CC for use in treatment of a NOVX related pathology. The antibodies and a
CC polypeptide having 95% sequence identity to NOVX polypeptide are useful
CC for treating a pathological state in a mammal. The present sequence
CC represents NOV8, a possible neuronal antigen-like protein
XX
SQ Sequence 351 AA;

Query Match 51.3%; Score 887.5; DB 4; Length 351;
Best Local Similarity 55.4%; Pred. No. 7.7e-82;
Matches 179; Conservative 53; Mismatches 84; Indels 7; Gaps 5;
QY 1 MAMTLLDWCRCGMDVNSQRTLLVWGPVNCDEAEIBETLQAAM-PQVSVMGLGRMFWRREE 59
Db 1 MTLRLLEDWCRCGMDVNSQRTLLVWGPVNCDEAEIBETLQAAM-PQVSVMGLGRMFWRREE 60
QY 60 NAKAALLEITGAVDYAAIPREMPKGGVWVLPKPTSDAEFLERLHLFLAREGWTVDV 119
Db 61 NRKVALVGLTAETSHALVPKEIPKGGIWRVIFKPPDPNTFLSRLNEFLAGEMTVGEL 120
QY 120 ARVLGFQNPPT--PTPG--PEMPAEMLNLYILDVNIQPLVESIWKLTFLSGKHPRAWRG 175
Db 121 SRAIGHENGSLDPEQCMIPEMWAPMLAQALE-ALQPALQCLKYKLVRSGRSPFEGEE 179
QY 176 NFDPMLEHTNEVLEEQVSDVEKRRRLMESLRGPAADVIRILKNNPAITTAECLEKALEQ 235
Db 180 EFGRWPFHTTQMIKAMQVDPVDEKRRRLLESRLRGPAADVIRILKNNPLITVDCECLQALEE 239
QY 236 VFGSVSSRDQAKFLNTYQNPCKLSAVIRLEPLLOKVEKGAIDKONVQARLEQVI 295
Db 240 VFGVTDNPRELQVKYLTYYQKDEKLSAVYLRLEPLLOKLVORGAIERDAVNOARLDQVI 299
QY 296 AGANHSGAIRRQLWLTGAGEGPG 318
Db 300 AGAVHK-TIRRELNPEDGPAFG 321
RESULT 6
AAB43023
ID AAB43023 standard; protein; 312 AA.
XX

DR WPI; 2000-468119/41.
 XX N-PSDB; AAA60837.
 PT Diagnosing paraneoplastic syndrome e.g. paraneoplastic limbic
 PT encephalitis or neoplasm e.g. colon cancer comprising assessing a test
 PT sample for the presence or absence of antibodies to a Ma family
 PT polypeptide.
 XX
 PS Claim 48; Fig 9-10; 27pp; Japanese.
 XX
 CC The present invention describes a method for diagnosing a paraneoplastic
 CC syndrome or neoplasm. The method comprises assessing a test sample for
 CC the presence or absence of antibodies to a Ma family polypeptide (I). The
 CC method is used to diagnose a paraneoplastic syndrome especially
 CC paraneoplastic limbic encephalitis and/or brainstem encephalitis or
 CC neoplasm in an individual. The method diagnoses the neoplasm by assessing
 CC antibodies to (I) preferably Ma1, which is indicative presence of breast
 CC cancer, colon cancer, parotid gland cancer, lung cancer, testicular
 CC cancer and germ-cell tumours or Ma2, which is indicative of testicular
 CC cancer, germ-cell tumour, and lung cancer. The present sequence is the
 CC Ma5 protein as given in the present invention
 XX
 SQ Sequence 463 AA;
 Query Match 44.3%; Score 766.5; DB 3; Length 463;
 Best Local Similarity 50.2%; Pred. No. 3e-69;
 Matches 157; Conservative 50; Mismatches 103; Indels 3; Gaps 2;
 QY 1 MAMTLLDWCRCMDVNSORTLLVWGPVNCDEABIEETLQAAAPQVS-YRMLGRMFWRREE 59
 DB 1 MPTLLQDWCRGEHNTTRCMLILGIPEDCGDEFEETLQACRHLGRYRVIGRMFRREE 60
 QY 60 NAKAALLELTGADVAAIAPREMPKGGYKVLFPKPTSDAEFLRLHLFLAREGTVQDV 119
 DB 61 NQAALLLELAQDIDYALLPREIPGKGGPWEIVKPRNSDGEFLNRLNRFLEERTVSDM 120
 QY 120 ARVLGFQNPPTPGPEMPAEMLY--ILDNVIQPLVESIWKRLTLFSGKGHPRAWGNF 177
 DB 121 NRVLGSDTNCAPRTVISPEFWTAAQLGAAVQLLEQMLRELRFVSGNTISIPGALAF 180
 QY 178 DPWLHTNEVLQVSDVEKRRRLMESLRGPAADVIRILKSNPAAITAECLKALEQVF 237
 DB 181 DAWLEHTTEMLQWQVPEGEKRRRLMECLRGPAQVWSGLRASNASITVEECLAALQOVF 240
 QY 238 GSVSSRDAQIKELNTYQNPCKLSAVYRLPELQKVKGAIDKNVQNAQLQEVLAG 297
 DB 241 GPFVSHKIAQVKLQKAYQGEAGEKVSFVIRLEPLQLQRAVENNVSRNNVQTRKRVLSG 300
 QY 298 ANHSGAIRQLWL 310
 DB 301 ATLPDKLRDKLKL 313
 RESULT 10
 AAB42315
 ID AAB42315 standard; protein; 463 AA.
 XX
 AC AAB42315;
 XX
 DT 08-FEB-2001 (first entry)
 XX
 DE Human ORFX ORF2079 polypeptide sequence SEQ ID NO:4158.
 XX
 KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
 KW vulnary; antiparotatic; antiparkinsonian; nootropic; neuroprotective;
 KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
 KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
 KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
 KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 KW cholesterol ester storage; systemic lupus erythematosus; infection;

KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
 KW thrombosis; contraceptive.
 OS Homo sapiens.
 XX
 PN WO200058473-A2.
 XX
 PD 05-OCT-2000.
 XX
 XX 31-MAR-2000; 2000WO-US008621.
 XX
 PR 31-MAR-1999; 99US-0127607P.
 PR 02-APR-1999; 99US-0127636P.
 PR 05-APR-1999; 99US-0127728P.
 PR 30-MAR-2000; 2000US-00540763.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Shinkets RA, Leach M;
 XX
 DR WPI; 2000-602362/57.
 DR N-PSDB; AAC76524.
 XX
 PT Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease.
 XX
 PS Claim 11; Page 3345-3347; 5507pp; English.
 XX
 CC AAC7446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytostatic; hepatotropic; vulnary;
 CC antiparotatic; antiparkinsonian; nootropic; neuroprotective; osteopathic;
 CC anticonvulsant; antiarthritic; immunosuppressant; immunostimulant;
 CC cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;
 CC dermatological; immunosuppressive; antiinflammatory; antibacterial;
 CC antiviral; antifungal; antirheumatic; antithyroid; and antianaemic. The
 CC sequences can be used for determining the presence of or predisposition
 CC to, or preventing or treating pathological conditions associated with an
 CC ORFX-associated disorder. The nucleic acids can be used to express ORFX
 CC proteins in gene therapy vectors. The proteins and nucleic acids may be
 CC used to treat cancers, proliferative disorders, neurodegenerative
 CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease,
 CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester
 CC storage, systemic lupus erythematosus, severe combined immunodeficiency
 CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune
 CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and
 CC cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to
 CC enhance coagulation; to inhibit thrombosis; and as a contraceptive
 XX
 SQ Sequence 463 AA;
 Query Match 44.3%; Score 766.5; DB 3; Length 463;
 Best Local Similarity 50.2%; Pred. No. 3e-69;
 Matches 157; Conservative 50; Mismatches 103; Indels 3; Gaps 2;
 QY 1 MAMTLLDWCRCMDVNSORTLLVWGPVNCDEABIEETLQAAAPQVS-YRMLGRMFWRREE 59
 DB 1 MPTLLQDWCRGEHNTTRCMLILGIPEDCGDEFEETLQACRHLGRYRVIGRMFRREE 60
 QY 60 NAKAALLELTGADVAAIAPREMPKGGYKVLFPKPTSDAEFLRLHLFLAREGTVQDV 119
 DB 61 NQAALLLELAQDIDYALLPREIPGKGGPWEIVKPRNSDGEFLNRLNRFLEERTVSDM 120
 QY 120 ARVLGFQNPPTPGPEMPAEMLY--ILDNVIQPLVESIWKRLTLFSGKGHPRAWGNF 177
 DB 121 NRVLGSDTNCAPRTVISPEFWTAAQLGAAVQLLEQMLRELRFVSGNTISIPGALAF 180
 QY 178 DPWLHTNEVLQVSDVEKRRRLMESLRGPAADVIRILKSNPAAITAECLKALEQVF 237
 DB 181 DAWLEHTTEMLQWQVPEGEKRRRLMECLRGPAQVWSGLRASNASITVEECLAALQOVF 240

QY 238 GSVSSRDAQIKFLNTYQNPGEKLSAYVIRLEPLQKVVEKGAIDKDNVNQARLEQVIAG 297
 DB 241 GPVESHKIAQVLCXKAYOAGEKVSFVLRLEPLQRAVENNVSRNVNQTLLKRVLSG 300
 QY 298 ANHSGAIRRQLWL 310
 DB 301 ATLDPKLRDKLKL 313
 RESULT 11
 ID AAO16179
 XX AAO16179 standard; protein; 452 AA.
 AC AAO16179;
 XX
 DT 28-MAR-2003 (first entry)
 XX
 DE Human protein #5.
 XX
 KW Human; vaccine; adult whole brain; foetal whole brain; tonsil;
 KW adult hippocampus; disease-associated SNP analysis; knockout mouse;
 KW disease model mouse; cancer; neurological disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200299103-A1.
 XX
 PD 12-DEC-2002.
 XX
 PF 27-MAY-2002; 2002WO-JP005134.
 XX
 PR 04-JUN-2001; 2001JP-00168370.
 PR 16-AUG-2001; 2001JP-00246915.
 XX
 XX (KAZU-) KAZUSA DNA RES INST FOUND.
 PA (PROT-) PROTEIN EXPRESS CO LTD.
 XX
 XX Ohara O, Nagase T, Nakajima D;
 PI
 DR MPI; 2003-140622/13.
 DR N-PSDB; AAL51207.
 XX
 XX DNA preferentially expressed in human adult and fetal brain tissue useful
 PT for diagnosis, treatment and analysis of cancer and mental disorders.
 XX
 PS Claim 1; Page 56-60; 73pp; Japanese.
 XX
 CC The invention comprises the amino acid and coding sequences of seven
 CC human proteins that are preferentially expressed in adult whole brain,
 CC foetal whole brain, tonsil and adult hippocampus tissue. The DNA
 CC sequences are useful for the analysis of disease-associated single
 CC nucleotide polymorphisms and the production of knockout and human disease
 CC model mice. The DNA and protein sequences of the invention are useful for
 CC the prevention (vaccine) and treatment of cancer and neurological
 CC disorders. The present amino acid sequence represents a human protein of
 CC the invention
 XX
 SQ Sequence 452 AA;
 Query Match 43.0%; Score 744; DB 6; Length 452;
 Best Local Similarity 46.0%; Pred. No. 5.8e-67;
 Matches 143; Conservative 64; Mismatches 100; Indels 4; Gaps 2;
 QY 1 MAMTLLEDWCRGMDVNSQRTLLVWGIPVNCDEAEIETLQAA-PQVSYRMLGRMFWRREE 59
 DB 5 MALTLLLEDWCKGMDMPKALLIVGIPMECSVEIQTIVKAGLQPLCAYRVLGRMFRED 64
 QY 60 NAKAALIELTGADVAAIPREMPGKGGVKVKLFKPTSDAEFLERLHLFLAREGWTVQDV 119
 DB 65 NAKAVFTELADVNTYTLPSHIPKGGSGWVVKPRNPDDFELSRNLNFKDEGRSMTDV 124
 QY 120 ARVLGFQNPFTTGPPEMPAEMNLNLDNVIQPLVESIWKRLTLFSGKGHPRAWRGNDP 179

DB 125 ARALGC---CSLPASLDAEVMQVRSPPLEPPKESMMYRKLKVFSGTASPSGGETTFD 181
 QY 180 WLEHTNEVLEEQVSDVEKRRLESLGPAADVIRILKSNPAITTAELCKALEQVFGS 239
 DB 182 WLEQVTEIMPIMQVSEVEKRRLESLGPAALSIWVLOANDSITVECLDALKQIFGD 241
 QY 240 VESSRDAQIKFLNTYQNPGEKLSAYVIRLEPLQKVVEKGAIDKDNVNQARLEQVIAG 299
 DB 242 KEDFRASQRFLOTSPKIGEKVSTFLLRLEPLQKAVHKSPLSVRSTDMIRLKLILARVA 301
 QY 300 HSGAIRRQLWL 310
 DB 302 MTPALRGKLEL 312
 RESULT 12
 ID ADA54410
 XX ADA54410 standard; protein; 399 AA.
 AC ADA54410;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 DE Human protein, SEQ ID 1978.
 XX
 KW Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;
 KW Gene Therapy; human; secretory protein; membrane proteins; cancer;
 KW inflammatory disease; osteoporosis; neurological disease.
 XX
 OS Homo sapiens.
 XX
 PN EP1293569-A2.
 XX
 PD 19-MAR-2003.
 XX
 PF 21-MAR-2002; 2002EP-00006586.
 XX
 PR 14-SEP-2001; 2001JP-00328381.
 PR 24-JAN-2002; 2002US-0350435P.
 XX
 XX (HELI-) HELIX RES INST.
 PA (REAS-) RES ASSOC BIOTECHNOLOGY.
 XX
 XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
 PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
 PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;
 XX
 DR MPI; 2003-395539/38.
 DR N-PSDB; ADA52771.
 XX
 XX New polynucleotides encoding full-length polypeptides, e.g. secretory
 PT and/or membrane proteins, useful for developing medicines for diseases in
 PT which the gene is involved, or as target molecules for gene therapy.
 XX
 PS Claim 14; SEQ ID NO 1978; 205pp; English.
 XX
 CC The present invention relates to novel human secretory or membrane
 CC proteins (ADA54072-ADA55710) and their coding sequences (ADA52433-
 CC ADA54071). The coding sequences are useful in the gene therapy of
 CC diseases caused by abnormalities of the proteins, e.g. cancer,
 CC inflammatory diseases, osteoporosis or neurological disease.
 XX
 SQ Sequence 399 AA;
 Query Match 35.9%; Score 620; DB 6; Length 399;
 Best Local Similarity 40.3%; Pred. No. 2.5e-54;
 Matches 137; Conservative 64; Mismatches 106; Indels 28; Gaps 6;
 QY 1 MAMTLLEDWCRGMDVNSQRTLLVWGIPVNCDEAEIETLQAA-PQVSYRMLGRMFWRREE 59
 DB 1 MAYTMLQDWCRWGVNARRLLILGIPEDCCDAEFQBSLEAALRPMGHFTVLGKAFRED 60

XX WPI; 2001-308781/32.
XX
XX New isolated nucleic acid molecule encoding a human secreted protein is
XX used in preventing, treating or ameliorating a medical condition.
XX
XX
XX Disclosure; Page 46; 519pp; English.
XX
XX AAD05121-AAD05203 represent cDNAs corresponding to 24 human secreted
XX protein genes, and AAE01232-AAE01311 represent the proteins they encode.
XX AAE01312-AAE01340 represent human secreted protein variants or fragments.
XX The secreted proteins and their genes are useful for preventing, treating
XX or ameliorating medical conditions, e.g., by protein or gene therapy.
XX Pathological conditions can be diagnosed by determining the amount of the
XX new protein in a sample or by determining the presence of mutations in
XX the new genes. Specific uses are described for each of the 24 genes,
XX based on the tissues in which they are most highly expressed, and include
XX developing products for the diagnosis or treatment of proliferative
XX disorders, cancer, tumours, foetal and developmental abnormalities,
XX haematopoietic disorders, diseases of the immune system, AIDS, autoimmune
XX diseases (e.g., rheumatoid arthritis), inflammation, allergies, disease),
XX neurological disorders (e.g., Alzheimer's disease, Parkinson's disease),
XX cognitive disorders, schizophrenia, asthma, skin disorders (e.g.,
XX psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders,
XX angiogenic disorders, kidney disorders, gastrointestinal disorders,
XX pregnancy-related disorders, endocrine disorders, and infectious. The
XX proteins can also be used to aid wound healing and epithelial cell
XX proliferation, to prevent skin aging due to sunburn, to maintain organs
XX before transplantation, for supporting cell culture of primary tissues,
XX to regenerate tissues, to identify their cognate ligands or binding
XX partners, and in chemotaxis, and can be used as a food additive or
XX preservative to modify storage properties. Antibodies specific for a
XX protein of the invention can be used in alleviating symptoms associated
XX with the disorders mentioned above, and in diagnostic immunoassays e.g.,
XX radioimmunoassay or enzyme linked immunosorbent assay (ELISA). The
XX present sequence represents a human secreted protein fragment referred to
XX in the disclosure of the invention
XX
XX Sequence 280 AA;
SQ
Query Match 33.0%; Score 570.5; DB 4; Length 280;
Best Local Similarity 54.9%; Pred. No. 1.7e-49;
Matches 123; Conservative 35; Mismatches 57; Indels 9; Gaps 5;
QY 99 AEFLEHLFLAREGWTQDVAVLGFQNPPT--PTPG--PEMAEMNLVYLDNVIOPLVE 154
DB 32 AEFGR--FLAGEGTVGLSALGHENGSLDPEQGMPEMWPAPMAQALE-ALQPALQ 87
QY 155 SIWYKELTLFGSGKHPRAWKGNFDPWLEHTNEVLEWQVSDVEKRRRLMESLRGPAADVI 214
DB 88 CLKYKLRVFGSGRESPEEGEEFGRWFFHTQMIKAWQVDFVEKRRRLLESRLGPAADVI 147
QY 215 RIKSNNPALTTRACLKALEQVFGSVSSRDAQIKFLNTYQNGEKLAVVLEPLQK 274
DB 148 RVLIKNPLITVDECLQALEEVEGVTDNPRELQVLYLTYYQDEKLSAYVLRLEPLQK 207
QY 275 VVEKGALDKNNVQARLEQVAGHSGAIRRRLWLTGAGEGPG 318
DB 208 LVORGAIERDAVNQARLDQVIAGAVHK-TIRRELNPEDGPAFG 250
RESULT 15
AAB12528
ID AAB12528 standard; protein; 283 AA.
XX
XX AAB12528;
XX
XX 02-NOV-2000 (first entry)
XX
XX Human Ma4 protein SEQ ID NO:11.
XX
XX Ma1; Ma2; Ma3; Ma4; Ma5; paraneoplastic limbic encephalitis; diagnosis;
XX paraneoplastic syndrome; neoplasm; colon cancer; brainstem encephalitis;
KW

KW breast cancer; parotid gland cancer; lung cancer; testicular cancer;
XX germ-cell tumour.
XX
XX Homo sapiens.
XX
XX JP2000146992-A.
XX
XX 26-MAY-2000.
XX
XX 10-NOV-1999; 99JP-00320171.
XX
XX 10-NOV-1999; 98US-00189527.
XX
XX (SLOK) SLOAN KETTERING INST CANCER RES.
XX
XX WPI: 2000-468119/41.
XX N-PSDB; AAE0836.
XX
XX Diagnosing paraneoplastic syndrome e.g. paraneoplastic limbic
XX encephalitis or neoplasm e.g. colon cancer comprising assessing a test
XX sample for the presence or absence of antibodies to a Ma family
XX polypeptide.
XX
XX Claim 48; Fig 7-8; 27pp; Japanese.
XX
XX The present invention describes a method for diagnosing a paraneoplastic
XX syndrome or neoplasm. The method comprises assessing a test sample for
XX the presence or absence of antibodies to a Ma family polypeptide (1). The
XX method is used to diagnose a paraneoplastic syndrome especially
XX paraneoplastic limbic encephalitis and/or brainstem encephalitis or
XX neoplasm in an individual. The method diagnoses the neoplasm by assessing
XX antibodies to (I) preferably Ma1, which is indicative presence of breast
XX cancer, colon cancer, parotid gland cancer, lung cancer, testicular
XX cancer and germ-cell tumours or Ma2, which is indicative of testicular
XX cancer, germ-cell tumour, and lung cancer. The present sequence is the
XX Ma4 protein as given in the present invention
XX
XX Sequence 283 AA;
SQ

Query Match 32.7%; Score 565; DB 3; Length 283;
Best Local Similarity 47.0%; Pred. No. 6.3e-49;
Matches 117; Conservative 47; Mismatches 73; Indels 12; Gaps 4;
QY 83 GKGGVWKVLFKPPPTSDAEFLERHLFLAREGWTQDVAVLGFQNPPTTP--GPEMPA 138
DB 3 GKGGVWKVLFKPPNQDTEFLERLNLFLKEGQTVSGMFRALGOEVSVPATVPCISPELLA 62
QY 139 EMLNYILDNVIOPLVESIWKRLTLFSGKHPRAWKGNFDPWLEHTNEVLEWQVSDVEK 198
DB 63 HLIGQAWAHAPQPLL-PMYRKLRFVFGSAVPAPEDSFEVWLEQATEIVKEWPVTEAEK 121
QY 199 RRLMESLRGPAADVIRILKSNNPALTTRACLKALEQVFGSVSSRDAQIKFLNTYQNGP 258
DB 122 KWLAEKSLGPAALDMHIVQADNPSISVECELEAFKQVFGSLSSRRTAQVRYLKYTEEG 181
QY 259 EKL SAYVIRLEPLQKQVKGALDKNNVQARLEQVAGHSGAIRRRLWLT--TGAGE 315
DB 182 EKVSAYVLRLETLTKRAVEKRAIPRRIADQVRLEQVMAGA----TLNQLWLCRLRELKQ 237
QY 316 GPGPKPLSV 324
DB 238 GPPPSFLEL 246

Search completed: September 21, 2004, 13:35:12
Job time : 63.7447 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 13:20:22 ; Search time 76.2367 Seconds
(without alignments)
1361.621 Million cell updates/sec

Title: US-10-037-860-4

Perfect score: 1729

Sequence: 1 MAMTLEDWCRGMDVNSQRT.....LTGAGEGPGPKPLSVAGADP 329

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_25.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phase.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1602	92.7	353	4	O95144 homo sapien
2	1602	92.7	353	4	Q8ng07 homo sapien
3	1525	88.2	353	11	Q8vHz4
4	1507	87.2	353	11	Q8vHz4 rattus norv
5	1501	86.8	353	11	Q9cyp2 mus musculu
6	887.5	51.3	351	4	Q8cic8 mus musculu
7	887.5	51.0	351	4	Q8cby2 homo sapien
8	875.5	50.6	351	6	Q8has1 homo sapien
9	870.5	50.3	352	11	Q95K14 macaca fasc
10	792.5	45.8	466	11	Q9erH6 mus musculu
11	776.5	44.9	364	6	Q8jzW8 mus musculu
12	770.5	44.6	364	4	Q9gmU3 macaca fasc
13	766.5	44.3	455	4	Q94959 homo sapien
14	766.5	44.3	463	4	Q9h044 homo sapien
15	750	43.4	192	4	Q9ul41 homo sapien
16	744	43.0	448	4	Q8nd90 homo sapien
					Q8net3 homo sapien

17	744	43.0	452	4	Q96PV4
18	704	40.7	365	11	Q8BHK0
19	620	35.9	399	4	Q96A40
20	564	32.6	283	4	Q9UL42
21	540	31.2	237	4	Q9H833
22	431	24.9	194	4	O95145
23	386.5	22.4	378	4	Q8N3H4
24	386.5	22.4	435	4	Q86V59
25	369.5	21.4	246	11	Q8C533
26	369.5	21.4	430	11	Q80VM8
27	335.5	19.4	403	4	Q8TE36
28	332.5	19.2	402	11	O9CZA5
29	332	19.2	402	4	Q8NLC1
30	315.5	18.2	393	11	Q9DBI7
31	315.5	18.2	393	11	Q8VD24
32	256.5	14.8	149	4	Q9UL43
33	194.5	11.2	327	11	Q8VC32
34	192	11.1	538	4	Q9H0W5
35	192	11.1	538	4	Q8TB26
36	114.5	6.6	2037	5	Q22511
37	111	6.4	555	16	Q82TR8
38	111	6.4	558	16	Q9JSW5
39	110.5	6.4	559	2	Q9F1V6
40	109.5	6.3	553	16	Q8PCN6
41	109	6.3	652	16	Q7V5L3
42	108	6.2	558	16	Q9K112
43	107	6.2	555	16	Q8XVB8
44	106.5	6.2	503	16	Q8NPK8
45	106	6.1	518	16	Q8G6F5

ALIGNMENTS

RESULT 1

O95144 ID O95144 PRELIMINARY; PRT; 353 AA.
AC O95144;
DT 01-MAY-1999 (TRENBLrel. 10, Created)
DT 01-OCT-2001 (TRENBLrel. 18, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Paraneoplastic neuronal antigen MAL.
GN MAL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN RP SEQUENCE FROM N.A.
RC TISSUE=Cerebellum;
RX MEDLINE=99158179; PubMed=10050892;
RA Dalmat J., Gultekin S.H., Voltz R., Hoard R., DesChamps T.,
RA Balmaceda C., Batchelor T., Gerstner E., Eichen J., Frennier J.,
RA Posner J.B., Rosenfeld M.R.;
RT "Mal, a novel neuron- and testis-specific protein, is recognized by
RT the serum of patients with paraneoplastic neurological disorders.";
RL Brain 122:27-39(1999).
[2]
RN RP SEQUENCE FROM N.A.
RC TISSUE=Cerebellum;
RX MEDLINE=99158179; PubMed=10050892;
RA Dalmat J., Rosenfeld M.R., Voltz R., Hoard R.,
RA Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
RL EMBL; AF037364; AAD13810.3; -;
DR Genew; HGNC:9158; FNMA1.
DR GO; GO:0005737; C:cytoplasm;
DR GO; GO:0005730; C:nucleolus; TAS.
DR GO; GO:0007417; P:central nervous system development; TAS.
DR GO; GO:0007283; P:spermatogenesis; TAS.
SQ SEQUENCE 353 AA; 39800 MW; 3BB41691AE89AD3D CRC64;

Query Match 92.7%; Score 1602; DB 4; Length 353;
Best Local Similarity 96.6%; Pred. No. 1.6e-120;
Matches 308; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

QY 1 MAMTLEDWCRGMDVNSQRTLLVWGIPVNCDAEAEIETLQAAMPQVSYRMGLGRMFWEEN 60
 DB 1 MAMTLEDWCRGMDVNSQRTLLVWGIPVNCDAEAEIETLQAAMPQVSYRMGLGRMFWEEN 60
 QY 61 AKAALELTGAVDYAAIPREMPGKGGVWVLPKPTSDAEFLERLHLFLAREGWTVDVA 120
 DB 61 AKAALELTGAVDYAAIPREMPGKGGVWVLPKPTSDAEFLERLHLFLAREGWTVDVA 120
 QY 121 RVLGFGNPTPTPGPEMPAEMLYILDNVIOPLVESIWKRLTLFSGKGHPRAWGNFDPW 180
 DB 121 RVLGFGNPTPTPGPEMPAEMLYILDNVIOPLVESIWKRLTLFSGKGHPRAWGNFDPW 180
 QY 181 LEHTNEVLEWQVSDVEKRRRLMESLRGPAADVIRILKSNPAITTAECCLKALEQVFGSV 240
 DB 181 LEHTNEVLEWQVSDVEKRRRLMESLRGPAADVIRILKSNPAITTAECCLKALEQVFGSV 240
 QY 241 ESSRDAQIKFLNTYQNPGEKLSAYVIRLEPLQKVVEKGAIDKDNVNQARLEQVIAGANH 300
 DB 241 ESSRDAQIKFLNTYQNPGEKLSAYVIRLEPLQKVVEKGAIDKDNVNQARLEQVIAGANH 300
 QY 301 SGAIRQLWLTGAGEGPP 319
 DB 301 SGAIRQLWLTGAGEGPP 319

RESULT 2

Q8NG07 ID Q8NG07 PRELIMINARY; PRT; 353 AA.
 AC Q8NG07;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Paraneoplastic antigen.
 GN PNMA1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP Schueller M.M., Jenne D.E., Schutze dit Belkner N., Hohlfeild R.,
 RA Voltz R.;
 RT "Mal (PNMA1).";
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Prostate;
 RA Strausberg R.;
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF320308; AAN05100.1; -;
 DR EMBL; BC039577; AAH39577.1; -;
 SQ SEQUENCE 353 AA; 39761 MW; EB7F5B6AEDA25961 CRC64;

Query Match 92.7%; Score 1602; DB 4; Length 353;
 Best Local Similarity 96.6%; Pred. No. 1.6e-120;
 Matches 308; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

QY 1 MAMTLEDWCRGMDVNSQRTLLVWGIPVNCDAEAEIETLQAAMPQVSYRMGLGRMFWEEN 60
 DB 1 MAMTLEDWCRGMDVNSQRTLLVWGIPVNCDAEAEIETLQAAMPQVSYRMGLGRMFWEEN 60
 QY 61 AKAALELTGAVDYAAIPREMPGKGGVWVLPKPTSDAEFLERLHLFLAREGWTVDVA 120
 DB 61 AKAALELTGAVDYAAIPREMPGKGGVWVLPKPTSDAEFLERLHLFLAREGWTVDVA 120
 QY 121 RVLGFGNPTPTPGPEMPAEMLYILDNVIOPLVESIWKRLTLFSGKGHPRAWGNFDPW 180
 DB 121 RVLGFGNPTPTPGPEMPAEMLYILDNVIOPLVESIWKRLTLFSGKGHPRAWGNFDPW 180
 QY 181 LEHTNEVLEWQVSDVEKRRRLMESLRGPAADVIRILKSNPAITTAECCLKALEQVFGSV 240
 DB 181 LEHTNEVLEWQVSDVEKRRRLMESLRGPAADVIRILKSNPAITTAECCLKALEQVFGSV 240

QY 241 ESSRDAQIKFLNTYQNPGEKLSAYVIRLEPLQKVVEKGAIDKDNVNQARLEQVIAGANH 300
 DB 241 ESSRDAQIKFLNTYQNPGEKLSAYVIRLEPLQKVVEKGAIDKDNVNQARLEQVIAGANH 300
 QY 301 SGAIRQLWLTGAGEGPP 319
 DB 301 SGAIRQLWLTGAGEGPP 319

RESULT 3

Q8VHZ4 ID Q8VHZ4 PRELIMINARY; PRT; 353 AA.
 AC Q8VHZ4;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Paraneoplastic onconeural protein MA1.
 GN MA1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Pellkofer H.L., Schubart A., Pagany M., Voltz R.D., Linington C.;
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF335505; AAL73196.1; -;
 SQ SEQUENCE 353 AA; 39793 MW; FFC3A717FDCEADA3 CRC64;

Query Match 88.2%; Score 1525; DB 11; Length 353;
 Best Local Similarity 90.0%; Pred. No. 2.5e-114;
 Matches 287; Conservative 15; Mismatches 17; Indels 0; Gaps 0;

QY 1 MAMTLEDWCRGMDVNSQRTLLVWGIPVNCDAEAEIETLQAAMPQVSYRMGLGRMFWEEN 60
 DB 1 MAMTLEDWCRGMDVNSQRTLLVWGIPVNCDAEAEIETLQAAMPQVSYRMGLGRMFWEEN 60
 QY 61 AKAALELTGAVDYAAIPREMPGKGGVWVLPKPTSDAEFLERLHLFLAREGWTVDVA 120
 DB 61 AKAALELTGAVDYAAIPREMPGKGGVWVLPKPTSDAEFLERLHLFLAREGWTVDVA 120
 QY 121 RVLGFGNPTPTPGPEMPAEMLYILDNVIOPLVESIWKRLTLFSGKGHPRAWGNFDPW 180
 DB 121 RVLGFGNPTPTPGPEMPAEMLYILDNVIOPLVESIWKRLTLFSGKGHPRAWGNFDPW 180
 QY 181 LEHTNEVLEWQVSDVEKRRRLMESLRGPAADVIRILKSNPAITTAECCLKALEQVFGSV 240
 DB 181 LEHTNEVLEWQVSDVEKRRRLMESLRGPAADVIRILKSNPAITTAECCLKALEQVFGSV 240
 QY 241 ESSRDAQIKFLNTYQNPGEKLSAYVIRLEPLQKVVEKGAIDKDNVNQARLEQVIAGANH 300
 DB 241 ESSRDAQIKFLNTYQNPGEKLSAYVIRLEPLQKVVEKGAIDKDNVNQARLEQVIAGANH 300
 QY 301 SGAIRQLWLTGAGEGPP 319
 DB 301 SGAIRQLWLTGAGEGPP 319

RESULT 4

Q9CYP2 ID Q9CYP2 PRELIMINARY; PRT; 353 AA.
 AC Q9CYP2;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE 5730402C15RIK protein.
 GN PNMA1 OR 5730402C15RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]

Db 121 SRALGHENSLDPEQGMIPENWAPMLAQALE-ALQPALQCLKYKKLVFSGRESPEPGEE 179
 QY 176 NPDPWLEHTNEVLEWQVSDVEKRRRLMESLRGPAADVIRILKSNPAITTAELKALEQ 235
 Db 180 EFGRMWFTTQMIKAWQVDPVDEKRRRLLESRLGPAADVIRILKSNPAITTAELKALEE 239
 QY 236 VFGSVSSDAIKFINTYQNPGEKLSAVVIRLEPLLQKVVEKGAIDKDNVNOARLEQVI 295
 Db 240 VFGVTDNPRELQVKYLTYYQDEKLSAVVIRLEPLLQKVVEKGAIDKDNVNOARLDQVI 299
 QY 296 AGANHSAGAIRRQLWLTGAGEGPG 318
 Db 300 AGAVHK-TIRRELNPEDGPAPG 321

RESULT 7

Q9HAS1 PRELIMINARY; PRT; 351 AA.
 AC Q9HAS1;
 DT 01-MAR-2001 (TReMBLrel. 16, Created)
 DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
 DE MAP-1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Cerebellum;
 RX MEDLINE=21264738; PubMed=11060313;
 RA Tan K.O., Tan K.M.L., Chan S.-L., Yee K.S.Y., Bevort M., Ang K.C.,
 RA Yu V.C.;
 RT "MAP-1, a novel proapoptotic protein containing a BH3-like motif that
 RT associates with Bax through its Bcl-2 homology domains.";
 RL J. Biol. Chem. 276:2802-2807(2001).
 DR EMBL; AF305550; AAG31786.1; -.
 SQ SEQUENCE 351 AA; 39521 MW; 8736401AD8A24EAD CRC64;

Query Match 51.0%; Score 882.5; DB 4; Length 351;
 Best Local Similarity 55.1%; Pred. No. 1.2e-62;
 Matches 178; Conservative 53; Mismatches 85; Indels 7; Gaps 5;

QY 1 MAMTLLDWCRCMDVNSQRTLLVWGIPVNCDAEIEETLQAAAM-POVSYRMILGRMFWREE 59
 Db 1 MTLRLLEDWCRCMDMNPRAKLLIAGISQSCSVAEIEEALQAGLAPLGEYRLILGRMPERDE 60
 QY 60 NAKAALLELTGADVAAIAPREMPGKGVWVLPKPTSDAEFLERHLFLAREGTVQDV 119
 Db 61 NRKVALVGLTAETSHALVPEIPGKGIWRVIFKPPDPDNTFLSRNEFLAGEGTVGEL 120
 QY 120 ARVLGFQNPPTPG--PMPAEMLYILDNVIOPLVESIWYKRLTLFSGKGHP 175
 Db 121 SRALGHENSLDPEQGMIPENWAPMLAQALE-ALQPALQCLKYKKLVFSGRESPEPGEE 179
 QY 176 NPDPWLEHTNEVLEWQVSDVEKRRRLMESLRGPAADVIRILKSNPAITTAELKALEQ 235
 Db 180 EFGRMWFTTQMIKAWQVDPVDEKRRRLLESRLGPAADVIRILKSNPAITTAELKALEE 239
 QY 236 VFGSVSSDAIKFINTYQNPGEKLSAVVIRLEPLLQKVVEKGAIDKDNVNOARLEQVI 295
 Db 240 VFGVTDNPRELQVKYLTYYQDEKLSAVVIRLEPLLQKVVEKGAIDKDNVNOARLDQVI 299
 QY 296 AGANHSAGAIRRQLWLTGAGEGPG 318
 Db 300 AGAVHK-TIRRELNPEDGPAPG 321

RESULT 8

Q95K14 PRELIMINARY; PRT; 351 AA.
 ID Q95K14
 AC Q95K14;
 DT 01-DEC-2001 (TReMBLrel. 19, Created)

DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
 DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
 DE Hypothetical protein.
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecinae; Macaca.
 OX NCBI_TaxID=9541;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Temporal Cortex;
 RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
 RA Suzuki Y., Sugano S., Hashimoto K.;
 RT "Isolation of full-length cDNA clones from macaque brain cDNA
 RT libraries.";
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB060854; BAB46873.1; -.
 KW Hypothetical protein.
 SQ SEQUENCE 351 AA; 39623 MW; C753054496A6FFB3 CRC64;

Query Match 50.6%; Score 875.5; DB 6; Length 351;
 Best Local Similarity 54.3%; Pred. No. 4.2e-62;
 Matches 178; Conservative 51; Mismatches 82; Indels 17; Gaps 5;

QY 1 MAMTLLDWCRCMDVNSQRTLLVWGIPVNCDAEIEETLQAAAM-POVSYRMILGRMFWREE 59
 Db 1 MTLRLLEDWCRCMDMNPRAKLLIAGISQSCSVAEIEEALQAGLAPLGEYRLILGRMPERDE 60
 QY 60 NAKAALLELTGADVAAIAPREMPGKGVWVLPKPTSDAEFLERHLFLAREGTVQDV 119
 Db 61 NRKVALVGLTAETSHALVPEIPGKGIWRVIFKPPDPDNTFLSRNEFLAGEGTVGEL 120
 QY 120 ARVLGFQNPPTPG--PMPAEMLYILDNVIOPLVESIWYKRLTLFSGKGHP 170
 Db 121 TRALAHEN-----GSLDLEQGMIPENWAPMLAQALE-ALQPALQCLKYKKLVFSGREP 174
 QY 171 RAWRGNFDPMLEHTNEVLEWQVSDVEKRRRLMESLRGPAADVIRILKSNPAITTAEL 230
 Db 175 EPGEEFGRWMPFTTQMIKAWQVDPVDEKRRRLLESRLGPAADVIRILKSNPAITTAEL 234
 QY 231 KALEQVFGSVSSDAIKFINTYQNPGEKLSAVVIRLEPLLQKVVEKGAIDKDNVNOAR 290
 Db 235 QALEEVFGVTDNPRELQVKYLTYYQDEKLSAVVIRLEPLLQKVVEKGAIDKDNVNOAR 294
 QY 291 LEQVIAGANHSAGAIRRQLWLTGAGEGPG 318
 Db 295 LDQVIAGAVHK-TIRRELNPEDGPAPG 321

RESULT 9

Q9ERH6 PRELIMINARY; PRT; 352 AA.
 ID Q9ERH6
 AC Q9ERH6;
 DT 01-MAR-2001 (TReMBLrel. 16, Created)
 DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
 DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
 DE MAP-1 (4930435G24RIK protein) (RIKEN CDNA 4930435G24 gene).
 GN MOAP1 OR 4930435G24RIK.
 OS Mus musculus (Mouse)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA Tan K.O., Fan K.M.L., Chan S.-L., Yee K.S.Y., Bevort M., Ang K.C.,
 RA Yu V.C.;
 RT "MAP-1: a novel pro-apoptotic protein containing a BH3-like motif that
 RT associates with Bax through its Bcl-2 homology domains.";
 RL J. Biol. Chem. 276:2802-2807(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Testis;

Query Match 44.9%; Score 776.5; DB 6; Length 364;
 Best Local Similarity 48.2%; Pred. No. 4e-54; Mismatches 64; Indels 13; Gaps 5;
 Matches 160; Conservative 64;

1 MAMTLEDWCRGMDVNSQRTLLVWGPVNCDAEIEETLQAAPOV-SYRMLGRMFWRRE 59
 1 MALALLEDCRIMSDEQKSLMVTGIPADFEAEIOEVLQETLKSIGSYRLGKIFRKE 60

60 NAKAALLELTGAVDYAAIPREMPGKGVWVLPKPTSDAEFLERHLFLAREGWTVDV 119
 61 NANAVLLELLEDTDSALPSEVQGGVWVIFKTPNQDTEFLERLNLFLKEGQTVSGM 120

120 ARVLGFQNPPTP-----GPMPAEMNLNDVIOPLVESIWKRLTLFSGKHGRAMRG 175
 121 FRALGEGSPATVPCISPPELLAHLLGQMAHAPQLL-PMYRKLRVFGSAVPAPEE 179

176 NPDPMLEHTNEVLEWQVSDVEKRRRLMESLGRPAADVIRILKSNPNPAITTAECLEKALQ 235
 180 PFEVWLEQATEIVKEMPVTEAEKKWLAESLRGPAALDLMHIVQADNPSISVEECLEAPKQ 239

236 VFGSVESRDAQIKFLNTYQNGEKLISAVIRLEPLQKVVEKGAIDKDNVQARLEQVI 295
 240 VFGSLESRTAQVRLKTYQEGEKVSAYVLELTLRLRAVEKRAIPRIADQVRLQV 299

296 AGANSHGAIRROLWL---TGAGEGPGPKPLSV 324
 300 AGA-----TLNQMLWCLRELKQDQPPPNFLEL 327

RESULT 12

Q94959 PRELIMINARY; PRT; 364 AA.

AC Q94959; (TREMBlrel. 10, Created)
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE Hypothetical protein KIAA0883 (Paraneoplastic associated brain-testis-cancer antigen).
 DE KIAA0883.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=9156230; PubMed=10048485;
 RA Nagase T., Ishikawa K., Suyama M., Kikuno R., Hirose M.,
 RA Miyajima N., Tanaka N., Kotani H., Nomura N., Ohara O.;
 RA "Prediction of the coding sequences of unidentified human genes. XII.
 RT The complete sequences of 100 new cDNA clones from brain which code
 RT for large proteins in vitro."
 RL DNA Res. 5:355-364 (1998).
 RP [2]
 RP SEQUENCE FROM N.A.
 RA Richey J.G., Dalmau J., Wade D., Rosenfeld M.R.;
 RT "Characterization of a Brain-Testis-Cancer Antigen."
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB020690; BAA74906.1; -
 DR EMBL; AF286487; AAG28165.1; -
 DR Genbank; U005162; PNM02.
 DR InterPro; IPR005162; Retrotrans_gag.
 DR Pfam; PF03732; Retrotrans_gag; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 364 AA; 41509 MW; 6E417AD96E3F0B93 CRC64;

Query Match 44.6%; Score 770.5; DB 4; Length 364;
 Best Local Similarity 47.6%; Pred. No. 1.2e-53;
 Matches 158; Conservative 65; Mismatches 96; Indels 13; Gaps 5;

1 MAMTLEDWCRGMDVNSQRTLLVWGPVNCDAEIEETLQAAPOV-SYRMLGRMFWRRE 59
 1 MALALLEDCRIMSDEQKSLMVTGIPADFEAEIOEVLQETLKSIGSYRLGKIFRKE 60

60 NAKAALLELTGAVDYAAIPREMPGKGVWVLPKPTSDAEFLERHLFLAREGWTVDV 119
 61 NANAVLLELLEDTDSALPSEVQGGVWVIFKTPNQDTEFLERLNLFLKEGQTVSGM 120

120 ARVLGFQNPPTP-----GPMPAEMNLNDVIOPLVESIWKRLTLFSGKHGRAMRG 175
 121 FRALGEGSPATVPCISPPELLAHLLGQMAHAPQLL-PMYRKLRVFGSAVPAPEE 179

176 NPDPMLEHTNEVLEWQVSDVEKRRRLMESLGRPAADVIRILKSNPNPAITTAECLEKALQ 235
 180 PFEVWLEQATEIVKEMPVTEAEKKWLAESLRGPAALDLMHIVQADNPSISVEECLEAPKQ 239

236 VFGSVESRDAQIKFLNTYQNGEKLISAVIRLEPLQKVVEKGAIDKDNVQARLEQVI 295
 240 VFGSLESRTAQVRLKTYQEGEKVSAYVLELTLRLRAVEKRAIPRIADQVRLQV 299

296 AGANSHGAIRROLWL---TGAGEGPGPKPLSV 324
 300 AGA-----TLNQMLWCLRELKQDQPPPNFLEL 327

Db 1 MALALLEDCRIMSDEQKSLMVTGIPADFEAEIOEVLQETLKSIGSYRLGKIFRKE 60

QY 60 NAKAALLELTGAVDYAAIPREMPGKGVWVLPKPTSDAEFLERHLFLAREGWTVDV 119
 61 NANAVLLELLEDTDSALPSEVQGGVWVIFKTPNQDTEFLERLNLFLKEGQTVSGM 120

QY 120 ARVLGFQNPPTP-----GPMPAEMNLNDVIOPLVESIWKRLTLFSGKHGRAMRG 175
 121 FRALGEGSPATVPCISPPELLAHLLGQMAHAPQLL-PMYRKLRVFGSAVPAPEE 179

QY 176 NPDPMLEHTNEVLEWQVSDVEKRRRLMESLGRPAADVIRILKSNPNPAITTAECLEKALQ 235
 180 PFEVWLEQATEIVKEMPVTEAEKKWLAESLRGPAALDLMHIVQADNPSISVEECLEAPKQ 239

QY 236 VFGSVESRDAQIKFLNTYQNGEKLISAVIRLEPLQKVVEKGAIDKDNVQARLEQVI 295
 240 VFGSLESRTAQVRLKTYQEGEKVSAYVLELTLRLRAVEKRAIPRIADQVRLQV 299

QY 296 AGANSHGAIRROLWL---TGAGEGPGPKPLSV 324
 300 AGA-----TLNQMLWCLRELKQDQPPPNFLEL 327

RESULT 13

Q9H0A4 PRELIMINARY; PRT; 455 AA.

AC Q9H0A4; (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Hypothetical protein.
 GN DKFZP434K225.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RX MEDLINE=21154917; PubMed=11230166;
 RA Wiemann S., Weil B., Wellenreuther R., Gassenhuber J., Glassl S.,
 RA Ansorge W., Boecker M., Blocker H., Bauersachs S., Blum H.,
 RA Lauber J., Duesterhoeft A., Beyer A., Koehrer K., Strack N.,
 RA Mewes H.W., Ottenwelder B., Obermaier B., Tampe J., Heubner D.,
 RA Wambutt R., Korn B., Klein M., Poustka A.;
 RT "Towards a Catalog of Human Genes and Proteins: Sequencing and
 RT Analysis of 500 Novel Complete Protein Coding Human cDNAs."
 RL Genome Res. 11:422-435 (2001).
 DR EMBL; ALI36878; CAB66812.1; -
 DR GO; GO:0003676; F:nucleic acid binding; IEA.
 DR InterPro; IPR001878; Znf_CCHC.
 DR Pfam; PF00098; zf-CCHC; 1.
 DR PRINTS; PR00939; C2HCZNFINGER.
 DR SMART; SM00343; Znf_C2HC; 1.
 DR PROSITE; PS0158; ZF_CCHC; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 455 AA; 51514 MW; 9AC8CC06685A8956 CRC64;

Query Match 44.3%; Score 766.5; DB 4; Length 455;
 Best Local Similarity 50.2%; Pred. No. 3.5e-53;
 Matches 157; Conservative 50; Mismatches 103; Indels 3; Gaps 2;

1 MAMTLEDWCRGMDVNSQRTLLVWGPVNCDAEIEETLQAAPOV-SYRMLGRMFWRRE 59
 1 MPTLLQDNCRCGRHLNTRCMLTILGIPEDCGEDEFETLQACRHLGRYVIGRMFRRE 60

60 NAKAALLELTGAVDYAAIPREMPGKGVWVLPKPTSDAEFLERHLFLAREGWTVDV 119
 61 NAAAILLEAQDIDYALLPREIPGKGPWEIVKPNPSDGEFLNRLNRFLEEBRTVSDM 120

120 ARVLGFQNPPTP-----GPMPAEMNLNDVIOPLVESIWKRLTLFSGKHGRAMRG 177
 121 NRVLGSDTNCSPAPRTVISPFEFTWAQTLGAAVQLLEQMLYRELRFVSGNTISIPGALAF 180

```
QY 178 DPWLEHTNEVLEEMQVSDVEKRRRLMESLRGPAADVIRILKSNPAITTAECLEKALQOVF 237
Db 181 DAWLEHTTEMLQWQVPEGEKRRRLMECLRGPAQVSLGRASNASITVEECALAQOVF 240
QY 238 GSVSSRDAQIKFLNTYONPGEKLSAYVIRLEPLLOKVVKEGAIDKDNVNQARLEQVVIAG 297
Db 241 GPVESHKIAQVKLCAYQAGEKVSFVLRLEPLLOQRAVENNVVSRNVTQLKRVLSG 300
QY 298 ANHSGAIRQIWL 310
Db 301 ATLPLDKLRDKLKL 313

RESULT 14
Q9UL41 PRELIMINARY; PRT; 463 AA.
AC Q9UL41
DT 01-MAY-2000 (TREMELrel. 13, Created)
DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Paraneoplastic neuronal antigen MA3.
GN MA3
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Gultekin S.H., Voltz R., Rosenfeld M.R., Gerstner E., Eichen J.,
RA Posner J.B., Dalmay J.;
RT "Identification of a novel cancer testis brain antigen using serum
RT antibodies from patients with testicular tumors and paraneoplastic
RT limbic encephalomyelitis.";
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF083116; AAF05627.1; -.
DR Genew; HGNC:18742; PNMA3.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR001878; Znf CCHC.
DR Pfam; PF00098; zf-CCHC; 1.
DR PRINTS; PR00939; C2HCZNFINGER.
DR SMART; SM00343; Znf C2HC; 1.
DR PROSITE; PS00158; ZF_CCHC; 1.
SQ SEQUENCE 463 AA; 52376 MW; 0843BF6601D8557C CRC64;

Query Match 44.3%; Score 766.5; DB 4; Length 463;
Best Local Similarity 50.2%; Pred. No. 3.6e-53;
Matches 157; Conservative 50; Mismatches 103; Indels 3; Gaps 2;

QY 1 MAMTLLEDWCRGMDVNSQRTLLVWGI PVNCDDEABIEETLQAAMPQVS- YRMILGRMFWRREE 59
Db 1 MPLTLQDWCRGHEHLNTRRCMLIIGIPEDCGEDFEETLQEACHLGRYVIGRMFRREE 60
QY 60 NAKAALLELTGADVAAIPREMPGKGVKVKFKPPTSDAEFLRLHLFLAREGWTQDV 119
Db 61 NAAQAILLELAQDIDYALLPREIPKGGPWEIVKPRNSDGEFLNRLNKRFLLEERRTVSDM 120
QY 120 ARVLGFQNPPTPGPEPAMWLY--ILDNVIQPLVESIWKRLTLESGKGHPRAWGNF 177
Db 121 NRVLGSDTNCAPRVTTISPFWTWAOQLGAQVLPLEQLYRELRFVSGNTISIPGALAF 180
QY 178 DPWLEHTNEVLEEMQVSDVEKRRRLMESLRGPAADVIRILKSNPAITTAECLEKALQOVF 237
Db 181 DAWLEHTTEMLQWQVPEGEKRRRLMECLRGPAQVSLGRASNASITVEECALAQOVF 240
QY 238 GSVSSRDAQIKFLNTYONPGEKLSAYVIRLEPLLOKVVKEGAIDKDNVNQARLEQVVIAG 297
Db 241 GPVESHKIAQVKLCAYQAGEKVSFVLRLEPLLOQRAVENNVVSRNVTQLKRVLSG 300
QY 298 ANHSGAIRQIWL 310
Db 301 ATLPLDKLRDKLKL 313
```

```
RESULT 15
Q8ND90 PRELIMINARY; PRT; 192 AA.
AC Q8ND90;
DT 01-OCT-2002 (TREMELrel. 22, Created)
DT 01-OCT-2002 (TREMELrel. 22, Last sequence update)
DT 01-OCT-2002 (TREMELrel. 22, Last annotation update)
DE Hypothetical protein (Fragment).
GN DKFZP434N171.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Poustka A., Wellenreuther R., Mewes H.W., Weil B., Wiemann S.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL834327; CAD38995.1; -.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 192 AA; 21456 MW; 5DB17CF6A760A628 CRC64;

Query Match 43.4%; Score 750; DB 4; Length 192;
Best Local Similarity 93.7%; Pred. No. 2.2e-52;
Matches 148; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 162 TLFSGKGHPRAWGNFDPWLEHTNEVLEEMQVSDVEKRRRLMESLRGPAADVIRILKSN 221
Db 1 TLFSGRDIPGPGERTFDPWLEHTNEVLEEMQVSDVEKRRRLMESLRGPAADVIRILKSN 60
QY 222 PAITTAECLEKALQOVFGSVSSRDAQIKFLNTYONPGEKLSAYVIRLEPLLOKVVKEGAI 281
Db 61 PAITTAECLEKALQOVFGSVSSRDAQIKFLNTYONPGEKLSAYVIRLEPLLOKVVKEGAI 120
QY 282 DKDNVQARLEQVITAGANHSGAIRRQLWLTGAGEGPGP 319
Db 121 DKDNVQARLEQVITAGANHSGAIRRQLWLTGAGEGPGP 158

Search completed: September 21, 2004, 13:39:30
Job time : 79.2367 secs
```

Blank

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 13:20:21; Search time 8.96902 Seconds
(without alignments)
1910.028 Million cell updates/sec

Title: US-10-037-860-4

Perfect score: 1729

Sequence: 1 MAMTLLEDWCRGMDVNSQRT.....LTGAGEGPGPKPLSVAGADP 329

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 141681 segs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1602	92.7	353	1	PMAL_HUMAN
2	1525	88.2	353	1	PMAL_RAT
3	1501	86.8	353	1	PMAL_MOUSE
4	887.5	51.3	351	1	MOP1_HUMAN
5	875.5	50.6	351	1	MOP1_MACFA
6	870.5	50.3	352	1	MOP1_MOUSE
7	776.5	44.9	364	1	PMAL2_MACFA
8	770.5	44.6	364	1	PMAL2_HUMAN
9	704	40.7	365	1	PMAL2_MOUSE
10	104.5	6.0	556	1	YJJK_HAEN
11	103	6.0	935	1	KINH_SYNEA
12	101	5.8	488	1	SYPA_AERPE
13	99.5	5.8	2564	1	SPCQ_HUMAN
14	98.5	5.7	424	1	FLAI_ARATH
15	98	5.7	513	1	Y26 MYCTU
16	97.5	5.6	759	1	PARC_CAUCR
17	96	5.6	468	1	EFLA_HYDAT
18	95.5	5.5	645	1	SYM_CLOPE
19	92.5	5.3	464	1	KIR4_YEAST
20	92.5	5.3	2431	1	POLN_SFV
21	91.5	5.3	378	1	B3GA_HUMAN
22	90	5.2	404	1	CPXJ_SACER
23	90	5.2	1660	1	VIT6_OSCBR
24	89	5.1	646	1	THIC_VIBPA
25	89	5.1	1755	1	PEPL_MOUSE
26	88.5	5.1	552	1	THSB_SULTO
27	88.5	5.1	808	1	SYFB_SYNP7
28	88	5.1	640	1	MYB_BOVIN
29	87	5.0	461	1	EFLA_APIPE
30	87	5.0	600	1	BRJ3_MOUSE
31	87	5.0	693	1	ERG_FUSNN
32	87	5.0	990	1	GOA2_HUMAN
33	87	5.0	6885	1	SNEZ2_HUMAN

34	86.5	5.0	605	1	VAA2_TREPA	O83541 treponema p
35	86.5	5.0	1155	1	IF2P_METJA	Q57710 methanococc
36	86.5	5.0	1364	1	KINI_CHICK	Q90631 gallus gall
37	86	5.0	316	1	FLGJ_SALTY	P15931 salmoneilla
38	86	5.0	364	1	GUNA_RUMAL	P23660 ruminococcu
39	86	5.0	424	1	PUR2_AQUAE	O66949 aquifex aeo
40	86	5.0	554	1	YJJK_ECOLI	P37797 escherichia
41	86	5.0	727	1	MUTB_PROFR	P11653 propionibac
42	85	4.9	926	1	RGP1_CAEEL	P34342 caenorhabdi
43	84.5	4.9	495	1	THSB_SULAC	Q9V2T4 mus musculu
44	84.5	4.9	546	1	DX5B_MOUSE	Q9D0R4 mus musculu
45	84.5	4.9	641	1	MYB_CHICK	P01103 gallus gall

ALIGNMENTS

RESULT 1

PMAL_HUMAN STANDARD; PRT; 353 AA.

AC Q8ND50; Q85144; Q8NG07;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Paraneoplastic antigen Mal (Neuron- and testis-specific protein 1)
DE (37 kDa neuronal protein).
GN PMAL OR MAL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RX MEDLINE=99158179; PubMed=10050892;
RC TISSUE=Cerebellum;
RA Dalmaz J., Gultekin S.H., Voltz R., Hoard R., DesChamps T.,
RA Balmaceda C., Batchelor T., Gerstner E., Eichen J., Premier J.,
RA Posner J.B., Rosenfeld M.R.;
RT "Mal, a novel neuron- and testis-specific protein, is recognized by
RT the serum of patients with paraneoplastic neurological disorders.";
RL Brain 122:27-39(1999).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,
RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez A.,
RA Fahey J., Helton E., Kettman M., Maman A., Rodriguez S., Sanchez A.,
RA Whiting M., Maman A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blackley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.B.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RP SEQUENCE OF 162-353 FROM N.A.

```
RC TISSUE=Testis;
RA Poustka A., Wellenreuther R., Mewes H.-W., Weill B., Wiemann S.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Nuclear; nucleolar. In tumor cells, it is
CC cytoplasmic.
CC -!- TISSUE SPECIFICITY: Testis and brain specific. In some patients
CC suffering from cancers, it is also specifically expressed by the
CC paraneoplastic tumor cells.
CC -!- MISCELLANEOUS: Antibodies against PNMA1 are present in sera from
CC patients suffering of paraneoplastic neurological disorders.
CC -!- SIMILARITY: Belongs to the PNMA family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF037364; RAD13810.3; -
CC EMBL; AF320308; AAN05100.1; -
CC EMBL; BC039577; AAH39577.1; -
CC EMBL; AL834327; CAD38995.1; -
CC Genew; HGNC:9158; PNMA1.
CC MIM; 604010; -
CC GO; GO:0005737; Cytoplasm; TAS.
CC GO; GO:0005730; C:nucleolus; TAS.
CC GO; GO:0007417; P:central nervous system development; TAS.
CC GO; GO:0007283; P:spermatogenesis; TAS.
CC Antigen; Tumor antigen; Nuclear protein.
CC DOMAIN 335 341 POLY-GLU.
CC SEQUENCE 353 AA; 39761 MW; EB7F5B6AEDA25961 CRC64;
CC -----
Query Match 92.7%; Score 1602; DB 1; Length 353;
Best Local Similarity 96.6%; Pred. No. 5.6e-119;
Matches 308; Conservative 1; Mismatches 10; Indels 0; Gaps 0;
CC -----
QY 1 MAMTLLDRCGMDVNSQRTLLVWGIPVNCDAEIEETLQAAMPQVSYRMLGRMFWEEN 60
DB 1 MAMTLLDRCGMDVNSQRTLLVWGIPVNCDAEIEETLQAAMPQVSYRMLGRMFWEEN 60
QY 61 AKAALELTGAVDYAAIPREMPGKGVWVKVLPKPPTSDAEFLERHLFLAREGWTVDVA 120
DB 61 AKAALELTGAVDYAAIPREMPGKGVWVKVLPKPPTSDAEFLERHLFLAREGWTVDVA 120
QY 121 RVLGFGNPTPTGPEMPAEMNLVLDNVIOPLVESIWKRLTLFSGKGHPRAWGNFDPW 180
DB 121 RVLGFGNPTPTGPEMPAEMNLVLDNVIOPLVESIWKRLTLFSGKGHPRAWGNFDPW 180
QY 181 LEHTNEVLEWQVSDVEKRRRLMESLRGPAADVIRILKSNNPATTTAECLKALEQVFGSV 240
DB 181 LEHTNEVLEWQVSDVEKRRRLMESLRGPAADVIRILKSNNPATTTAECLKALEQVFGSV 240
QY 241 ESSRDAQIKFLNTYQNPGEKLSAYVIRLEPFLQKVVKEGAIDKDNVNQARLEQVIAGNH 300
DB 241 ESSRDAQIKFLNTYQNPGEKLSAYVIRLEPFLQKVVKEGAIDKDNVNQARLEQVIAGNH 300
QY 301 SGAIRQLMTGAGEGPGP 319
DB 301 SGAIRQLMTGAGEGPGP 319
CC -----
RESULT 2
PNMA1_RAT ID PNMA1_RAT STANDARD; PRT; 353 AA.
AC Q8VHZ4;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DE Paraneoplastic antigen Mal homolog.
GN PNMA1 OR MAL.
OS Rattus norvegicus (Rat).
```

```
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RX MEDLINE=99158179; PubMed=10050892;
RA Dalmau J., Gultekin S.H., Voltz R., Hoard R., DesChamps T.,
RA Balmaceda C., Batchelor T., Gerstner E., Eichen J., Frenniet J.,
RA Posner J.B., Rosenfeld M.R.;
RT 'Mal, a novel neuron- and testis-specific protein, is recognized by
RT the serum of patients with paraneoplastic neurological disorders.';
RL Brain 122:27-39(1999).
CC -!- SUBCELLULAR LOCATION: Nuclear; nucleolar (By similarity).
CC -!- TISSUE SPECIFICITY: Testis and brain specific.
CC -!- SIMILARITY: Belongs to the PNMA family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF335505; AAL73196.1; -
CC Nuclear protein.
CC DOMAIN 335 341 POLY-GLU.
CC SEQUENCE 353 AA; 39793 MW; FFC3A717FDCEDA3 CRC64;
CC -----
Query Match 88.2%; Score 1525; DB 1; Length 353;
Best Local Similarity 90.0%; Pred. No. 6.6e-113;
Matches 287; Conservative 15; Mismatches 17; Indels 0; Gaps 0;
CC -----
QY 1 MAMTLLDRCGMDVNSQRTLLVWGIPVNCDAEIEETLQAAMPQVSYRMLGRMFWEEN 60
DB 1 MAMTLLDRCGMDVNSQRTLLVWGIPVNCDAEIEETLQAAMPQVSYRMLGRMFWEEN 60
QY 61 AKAALELTGAVDYAAIPREMPGKGVWVKVLPKPPTSDAEFLERHLFLAREGWTVDVA 120
DB 61 AKAALELTGAVDYAAIPREMPGKGVWVKVLPKPPTSDAEFLERHLFLAREGWTVDVA 120
QY 121 RVLGFGNPTPTGPEMPAEMNLVLDNVIOPLVESIWKRLTLFSGKGHPRAWGNFDPW 180
DB 121 RVLGFGNPTPTGPEMPAEMNLVLDNVIOPLVESIWKRLTLFSGKGHPRAWGNFDPW 180
QY 181 LEHTNEVLEWQVSDVEKRRRLMESLRGPAADVIRILKSNNPATTTAECLKALEQVFGSV 240
DB 181 LEHTNEVLEWQVSDVEKRRRLMESLRGPAADVIRILKSNNPATTTAECLKALEQVFGSV 240
QY 241 ESSRDAQIKFLNTYQNPGEKLSAYVIRLEPFLQKVVKEGAIDKDNVNQARLEQVIAGNH 300
DB 241 ESSRDAQIKFLNTYQNPGEKLSAYVIRLEPFLQKVVKEGAIDKDNVNQARLEQVIAGNH 300
QY 301 SGAIRQLMTGAGEGPGP 319
DB 301 SGAIRQLMTGAGEGPGP 319
CC -----
RESULT 3
PNMA1_MOUSE ID PNMA1_MOUSE STANDARD; PRT; 353 AA.
AC Q8C1C8; Q9CYP2;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DE Paraneoplastic antigen Mal homolog.
GN PNMA1 OR MAL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
```


RC STRAIN=C57BL/6J; TISSUE=Brain, and Embryo;
 RX MEDLINE=22354683; PubMed=12466851;
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
 RA Nikaide I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
 RA Iagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
 RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
 RA Blake J.A., Bradt D., Brusci V., Chothia C., Corbani L.E., Cousins S.,
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
 RA Grimmond S., Guscinich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
 RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
 RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertele G., Pesole G.,
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
 RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
 RA Verardo R., Wagner L., Wahlstedt C., Wang Y., Watanabe Y., Wells C.,
 RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
 RA Yuan Z., Zavalan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
 RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
 RA Birney E., Hayashizaki Y.;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573 (2002).
 CC -!- SIMILARITY: Belongs to the PNMA family.
 CC -!- SIMILARITY: Belongs to the PNMA family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AK017476; BAB30762.1; -;
 DR EMBL; AK028331; BAC25885.1; -;
 DR MGD; MGI:2180564; Pnmal.
 KW Nuclear protein.
 FT DOMAIN 336 341 POLY-GLU.
 FT CONFLICT 136 136 T -> M (IN REF. 1; BAB30762).
 SQ SEQUENCE 353 AA; 39688 MW; 18CEDC3AC4E70939 CRC64;
 Query Match 86.8%; Score 1501; DB 1; Length 353;
 Best Local Similarity 88.7%; Pred. No. 5.1e-111;
 Matches 283; Conservative 15; Mismatches 21; Indels 0; Gaps 0;
 Qy 1 MAMTLLDWCRCMDVNSQRTLLVWGIPVNCDEAEIEETLQAMPQVSYRMLGRMFWREN 60
 Db 1 MAMTLLDWCRCMDVNSQRTLLVWGIPVNCDEAEIEETLQAMPQVSYRMLGRMFWREN 60
 Qy 61 AKAAELLELTGADVAAAPREMPGKGVVWVFKPPTSDAEFLERHLHFLAREGWTVDQVA 120
 Db 61 AKAAELLELTGADVDSLIPREMPGKGLVWVFKPPTSDAEFLERHLHFLAREGWTVDQVA 120
 Qy 121 RVLGQNPPTPGMPAEMLYILDNVIQPLVESIWYKRLTLPSGKHPRAWRGNEPDW 180
 Db 121 RVLGQNPAPAGPPTPAEMLYILDNVIQPLVESIWYKRLTLPSGKHPRAWRGNEPDW 180
 Qy 181 LEHTNEVLEEQVSDVEKRRRLMESLRGPAADVIRILKSNPNATTTAECLKALEQVFGSV 240
 Db 181 LEHNSNIEEQVSDVEKRRRLMESLRGPAADVIRILKSNPNATTTAECLKALEQVFGSV 240
 Qy 241 ESSRDAQIKFNTYQNPGEKLSAVYIRLEPLQKVKVEKAIDKDNVQARLEQVIAGNH 300
 Db 241 ESSRDAQVRFNTYQNPGEKLSAVYIRLEPLQKVKVEKAIDKDNVQARLEQVIAGNH 300

Qy 301 SCATRRQLWLTGAGEGPGP 319
 Db 301 SCALRRQLWLTGAGEGPGP 319
 RESULT 4
 MOPI_HUMAN
 ID MOPI_HUMAN STANDARD; PRT; 351 AA.
 AC Q96BY2; Q9H833; Q9HAS1;
 DT 15-MAR-2004 (Rel. 43, Created)
 DT 15-MAR-2004 (Rel. 43, Last sequence update)
 DE Modulator of apoptosis 1 (MAP-1) (Paraneoplastic antigen MA4).
 GN MOAP1 OR PNMA4.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., CHARACTERIZATION, AND MUTAGENESIS OF LEU-120;
 RP 120-LEU-ARG-127 AND 125-GLY--GLU-127.
 RC TISSUE=Cerebellum;
 RX MEDLINE=21264738; PubMed=11060313;
 RA Tan K.O., Tan K.M.L., Chan S.-L., Yee K.S.Y., Bevtort M., Ang K.C.,
 RA Yu V.C.;
 RT "MAP-1, a novel proapoptotic protein containing a BH3-like motif that
 RT associates with Bax through its Bcl-2 homology domains.";
 RL J. Biol. Chem. 276:2802-2807 (2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Retinoblastoma;
 RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
 RA Nishikawa T., Nagai K., Sato H., Sugano S., Shiratori A., Sudo H.,
 RA Wagatsuma M., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
 RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
 RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
 RA Yanamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,
 RA Ninomiya K., Iwayanagi T.;
 RT "NEO human cDNA sequencing project.";
 RT Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Skin;
 RX MEDLINE=22386257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udén T.B., Toshyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.B.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 CC -!- FUNCTION: Mediates caspase-dependent apoptosis.
 CC -!- SUBUNIT: Homodimer. Binds BAX, BCL2 and BCLX.
 CC -!- TISSUE SPECIFICITY: Ubiquitously expressed with higher levels in
 CC heart and brain.
 CC -!- DOMAIN: The BH3-like domain is required for association with BAX
 CC and for mediating apoptosis. The three BH domains (BH1, BH2, and
 CC BH3) of BAX are all required for mediating protein-protein
 CC interaction.

CC -!- SIMILARITY: Belongs to the PNMA family.
CC -!- CAUTION: Ref.2 sequence differs from that shown due to a
CC frameshift in position 102.
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF305550; BAB1786.1; -;
CC EMBL; AK024029; BAB14788.1; ALT_SEQ.
CC EMBL; BC015044; AAH15044.1; -;
CC Genew; HGNC:16658; MOAP1.
KW Apoptosis.
FT SITE 120 127 BH3-LIKE.
FT MUTAGEN 120 127
FT MUTAGEN 120 127
FT MUTAGEN 125 127
FT MUTAGEN 244 244
FT CONFLICT 258 258 T -> A (IN REF. 2).
FT CONFLICT 259 259 Y -> H (IN REF. 2).
FT CONFLICT 259 259 Q -> H (IN REF. 1).
SQ SEQUENCE 351 AA; 39512 MW; 5310142AC02B563C CRC64;

Query Match 51.3%; Score 887.5; DB 1; Length 351;
Best Local Similarity 55.4%; Pred. No. 1.2e-62;
Matches 179; Conservative 53; Mismatches 84; Indels 7; Gaps 5;

QY 1 MAMTLLDWCRCMDVNSQRTLLVWGIPIVNCDAEIEETLQAAM-POVSYRMGLRMFWREE 59
DB 1 MTRLLDWCRCMDMNPRAKALLTAGISQSCVAIEEALQAGLAPGEYRLGLRMFRRDE 60

QY 60 NAKAALLELTGAVDYAAIPREMPGKGWVKVLPKPTSDAEFLERLHLFLAREGTVQDV 119
DB 61 NRKVALVGLTAETSHALVPEKIPGKGIWRVIFKPPDPNTFLSRNLEFLAGEGTVGEL 120

QY 120 ARVLGQNPPT-PTPG--PEMPAEMLYILDNVIQPLVESIWKYKLTFLSGKHPRAWRG 175
DB 121 SRALCHENGSLDPEQGMIPENWAPMLAQALE-ALQALQCLYKLVKLVFSGRSPGEE 179

QY 176 NEDPWLHNTNEVLEWQVSDVEKRRRLMESLRGPAADVIRILKSNPAITTAECLEQ 235
DB 180 EFGRMFHTTQMIKAWQVDPVEKRRRLLESRLRGFALDVIRVLKINNPLITVDECLQALEE 239

QY 236 VFGSVSSRDAQIKFLNTYQNPGEKLSAYVIRLEPLLOKVEKGAIDKDNVNOARLEQVI 295
DB 240 VFGVTNPRELQVKYLTYYQKDEEKL SAYVIRLEPLLOKLVQVGAIERDAVNQARLDQVI 299

QY 296 AGANHSAGAIRQLWLTGAGEGPG 318
DB 300 AGAVHK-TIRRELNPEDGPAPG 321

RESULT 5
MOPL_MACFA STANDARD; PRT; 351 AA.
ID MOPL_MACFA
AC Q95K14;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Modulator of apoptosis 1 (MAP-1).
GN MOAP1.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]

RP SEQUENCE FROM N.A.
RC TISSUE=Temporal Cortex;
RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
RA Suzuki Y., Sugano S., Hashimoto K.;
RT Isolation of full-length cDNA clones from macaque brain cDNA
RT libraries";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Mediates caspase-dependent apoptosis (By similarity).
CC -!- SUBUNIT: Homodimer. Binds BAX, BCL2 and BCLX (By similarity).
CC -!- DOMAIN: The BH3-like domain is required for association with BAX
CC and for mediating apoptosis. The three BH domains (BH1, BH2, and
CC BH3) of BAX are all required for mediating protein-protein
CC interaction (By similarity).
CC -!- SIMILARITY: Belongs to the PNMA family.
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB060854; BAB46873.1; -;
KW Apoptosis.
FT SITE 120 127 BH3-LIKE.
FT DOMAIN 336 339 POLY-GLU.
SQ SEQUENCE 351 AA; 39623 MW; C7530E4496A6FB3 CRC64;

Query Match 50.6%; Score 875.5; DB 1; Length 351;
Best Local Similarity 54.3%; Pred. No. 1.1e-61;
Matches 178; Conservative 51; Mismatches 82; Indels 17; Gaps 5;

QY 1 MAMTLLDWCRCMDVNSQRTLLVWGIPIVNCDAEIEETLQAAM-POVSYRMGLRMFWREE 59
DB 1 MTRLLDWCRCMDMNPRAKALLTAGISQSCVAIEEALQAGLAPGEYRLGLRMFRRDE 60

QY 60 NAKAALLELTGAVDYAAIPREMPGKGWVKVLPKPTSDAEFLERLHLFLAREGTVQDV 119
DB 61 NRKVALVGLTAETSHALVPEKIPGKGIWRVIFKPPDSNTFLSRNLEFLAGEGTVGEL 120

QY 120 ARVLGQNPPTPG-----PEMPAEMLYILDNVIQPLVESIWKYKLTFLSGKHGP 170
DB 121 TRALAHEN-----GSLDLQGMIPENWAPMLAQALE-ALQALQCLYKLVKLVFSGRPP 174

QY 171 RAWRGNFDPLWLEHTNEVLEWQVSDVEKRRRLMESLRGPAADVIRILKSNPAITTAECLE 230
DB 175 EFGEEFEGRMFHTTQMIKAWQVDPVEKRRRLLESRLRGFALDVIRVLKINNPLITVDECL 234

QY 231 KALEQVFGSVSSRDAQIKFLNTYQNPGEKLSAYVIRLEPLLOKVEKGAIDKDNVNOAR 290
DB 235 QALEEVFGVTNPRELQVKYLTYYQKDEEKL SAYVIRLEPLLOKLVQVGAIERDAVNQAR 294

QY 291 LEQVIAGANHSAGAIRQLWLTGAGEGPG 318
DB 295 LDQVIAGVHK-TIRRELNPEDGPAPG 321

RESULT 6
MOPL_MOUSE STANDARD; PRT; 352 AA.
ID MOPL_MOUSE
AC Q9ERH6;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Modulator of apoptosis 1 (MAP-1).
GN MOAP1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.

FT DOMAIN 333 338 POLY-GLU.
 SQ SEQUENCE 364 AA; 41350 MW; 0CF72210D7EC1524 CRC64;
 Query Match 44.9%; Score 776.5; DB 1; Length 364;
 Best Local Similarity 48.2%; Pred. No. 7.1e-54;
 Matches 160; Conservative 64; Mismatches 95; Indels 13; Gaps 5;
 QY 1 MAMTLLDWCRCMDVNSQRTLLVWGIPVNCDSAEIETLQAAMPQV-SYRMGLRMFWREE 59
 DB 1 MALALLEDCRIMSVDEQSLMTVGIPVDEAEIOEVLQETLKSGLSYLLGKIFRQE 60
 QY 60 NAKAALLELTGAVDYAAIPREMPGKGVWVKVLFKPTSDAEFLERHLFLAREGTVQDV 119
 DB 61 NANAVILLELLETDVSAIPSEVQGGVWVKVFKPTNQDTEFLERLNLFLKGGQTVSGM 120
 QY 120 ARVLGFQNPPTTP---GPEMAELNYILDVNIQPLVESIMYKRLTLFSGKHPRWRG 175
 DB 121 FRALGHEGVSPTVPCISPELLAHLGQAMAHAPQLL-PMRYRKURVFGSVAVPAPEE 179
 QY 176 NFDPMLEHTNEVLEEQVSDVKRRRLMESLRGPAADVIRILKSNPNPAITTAECLEKALQ 235
 DB 180 PREVMLEQATEIVKEWPTVEAEKKRWLAESLRGPAALDLMHIVQADNPSISVEECLEAFKQ 239
 QY 236 VFGSVESSDAQIKFLNTYONPGEKLSAYVIRLEPLLQKVVEKAGDKDNVNOARLEQVI 295
 DB 240 VFGSLESRRTAQVRYLKYTQEEGKVSAYVIRLETLRRRAVEKRAIPRRADQVRLEQVM 299
 QY 296 AGANHSAGAIRROLWL---TCAGEGPGPKPLSV 324
 DB 300 AGA-----TLNQMLWCRLRELKQGGPPNFLEL 327

RESULT 8

PMA2_HUMAN STANDARD; PRT; 364 AA.
 AC Q9UL42; Q95145; Q9UL43;
 DT 15-MAR-2004 (Rel. 43, Created)
 DT 15-MAR-2004 (Rel. 43, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Paraneoplastic antigen M2 (Onconeural antigen M2) (Paraneoplastic
 DE neuronal antigen M2) (40 kDa neuronal protein).
 GN PMA2 OR M2 OR KIA00883.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
 RC TISSUE=Cerebellum;
 RX MEDLINE=99270611; PubMed=10362822;
 RA Voltz R., Gultekin S.H., Rosenfeld M.R., Gerstner E., Eichen J.,
 RA Posner J.B., Dalmay J.;
 RT "A serologic marker of paraneoplastic limbic and brain-stem
 RT encephalitis in patients with testicular cancer."
 RL New Engl. J. Med. 340:1788-1795 (1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=99156230; PubMed=10048495;
 RA Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirose M.,
 RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. XII.
 RT The complete sequences of 100 new cDNA clones from brain which code
 RT for large proteins in vitro."
 RL DNA Res. 5:355-364 (1998).
 RN [3]
 RP IDENTIFICATION, AND SUBCELLULAR LOCATION.
 RX MEDLINE=99158179; PubMed=10050892;
 RA Dalmay J., Gultekin S.H., Voltz R., Hoard R., DesChamps T.,
 RA Balmaceda C., Batchelor T., Gerstner E., Eichen J., Frenniet J.,
 RA Posner J.B., Rosenfeld M.R.;
 RT "M2, a novel neuron- and testis-specific protein, is recognized by
 RT the serum of patients with paraneoplastic neurological disorders."

RL Brain 122:27-39 (1999).
 CC -!- SUBCELLULAR LOCATION: Nuclear; nucleolar.
 CC -!- TISSUE SPECIFICITY: Brain specific. In some patients suffering
 CC from cancers, it is also specifically expressed by the testicular
 CC tumor cells.
 CC -!- MISCELLANEOUS: Antibodies against PMA2 are present in sera from
 CC patients suffering of paraneoplastic neurological disorders.
 CC -!- SIMILARITY: Belongs to the PNMA family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; AF037365; AAD02098.1; -;
 CC EMBL; AF083114; AAF05625.1; -;
 CC EMBL; AF083115; AAF05626.1; -;
 CC EMBL; AF286487; AAG28165.1; -;
 CC EMBL; AB020690; BAA74906.1; -;
 CC Genew; HGNC:9159; PNMA2.
 CC MIM; 603970; -;
 CC DR Antigen; Tumor antigen; Nuclear protein.
 KW DOMAIN 333 338 POLY-GLU.
 FT CONFLICT 128 129 GV -> AL (IN REF. 1; AAD02098).
 FT CONFLICT 141 141 L -> I (IN REF. 1; AAD02098).
 FT CONFLICT 257 257 T -> P (IN REF. 1; AAF05625).
 FT CONFLICT 278 278 R -> K (IN REF. 1; AAF05626).
 SQ SEQUENCE 364 AA; 41509 MW; 6E417AD9683F0E93 CRC64;

Query Match 44.6%; Score 770.5; DB 1; Length 364;

Best Local Similarity 47.6%; Pred. No. 2.1e-53;
 Matches 158; Conservative 65; Mismatches 96; Indels 13; Gaps 5;
 QY 1 MAMTLLDWCRCMDVNSQRTLLVWGIPVNCDSAEIETLQAAMPQVS-YRMGLRMFWREE 59
 DB 1 MALALLEDCRIMSVDEQSLMTVGIPADFEAEIOEVLQETLKSGLRYLLGKIFRQE 60
 QY 60 NAKAALLELTGAVDYAAIPREMPGKGVWVKVLFKPTSDAEFLERHLFLAREGTVQDV 119
 DB 61 NANAVILLELLETDVSAIPSEVQGGVWVKVFKPTNQDTEFLERLNLFLKGGQTVSGM 120
 QY 120 ARVLGFQNPPTTP---GPEMAELNYILDVNIQPLVESIMYKRLTLFSGKHPRWRG 175
 DB 121 FRALGHEGVSPTVPCISPELLAHLGQAMAHAPQLL-PMRYRKURVFGSVAVPAPEE 179
 QY 176 NFDPMLEHTNEVLEEQVSDVKRRRLMESLRGPAADVIRILKSNPNPAITTAECLEKALQ 235
 DB 180 SPEVMLEQATEIVKEWPTVEAEKKRWLAESLRGPAALDLMHIVQADNPSISVEECLEAFKQ 239
 QY 236 VFGSVESSDAQIKFLNTYONPGEKLSAYVIRLEPLLQKVVEKAGDKDNVNOARLEQVI 295
 DB 240 VFGSLESRRTAQVRYLKYTQEEGKVSAYVIRLETLRRRAVEKRAIPRRADQVRLEQVM 299
 QY 296 AGANHSAGAIRROLWL---TCAGEGPGPKPLSV 324
 DB 300 AGA-----TLNQMLWCRLRELKQGGPPNFLEL 327

RESULT 9

PMA2_MOUSE
 ID_PMA2_MOUSE STANDARD; PRT; 365 AA.
 AC Q8BHK0;
 DT 15-MAR-2004 (Rel. 43, Created)
 DT 15-MAR-2004 (Rel. 43, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Paraneoplastic antigen M2 homolog.
 GN PNMA2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

```

OX NCBI_TaxID=10090;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Brain cortex;
RX MEDLINE=22354683; PubMed=12466851;
RA Okazaki Y., Furuno M., Kasukawa T., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Nikaide I., Oeato N., Saito R., Saito R., Saito R., Schonbach C., Gojobori T.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusci V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Hiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
CC -!- SUBCELLULAR LOCATION: Nuclear; nucleolar (By similarity).
CC -!- SIMILARITY: Belongs to the PMA family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
CC EMBL; AK043718; BAC31626.1; -
CC DR EMBL; AK043910; BAC31700.1; -
CC DR MGD; MGI:2444129; Pma2.
CC KW Nuclear protein.
CC FT DOMAIN 333 338 POLY-GLU.
CC SQ SEQUENCE 365 AA; 41201 MW; 0F90C940B9D843D9 CRC64;
Query Match
Best Local Similarity 40.7%; Score 704; DB 1; Length 365;
Matches 144; Conservative 65; Mismatches 95; Indels 10; Gaps 5;
QY 1 MAMTLLDWCRCGMDVNSORTLLVWGIPVNCDEAIEETLQAMPQV-SYRMLGRMPWREE 59
Db 1 MAVALLBWKIKMGVDVQKSLVVDVDPVDCGEPIQTVLQELKCVGSYRLLGKIFQKD 60
QY 60 NAKAALLETGAVDYAIPREMPGKGGVKKVLFKPTSDAEFLRLHLFLAREGWTQDV 119
Db 61 NTSVVLVELMEDTMSVSPSEVQKGGVKKVIFKPTQDTEFLQRLNLFLEKSGQTWAGM 120
QY 120 ARVLGFQNPPT-TP---GPMPAEMLNLYLDNVIQPLVESIWKRLTLFGKGHPRAMRG 175
Db 121 FRALKHGVSPATPCTSPPELLAHLTGQAMVHGQRELL-PVKYCKMRIFSGSTAAPEEE 179
QY 176 NFDPLWLEHTNLEWQVSDVEKRRRLMESLRGPAADVIRILKSNPAITTAECIKALEQ 235
Db 180 PFEVWLQATEIAKEWPIPEAKRKRWAESLRGPAALDMHIVQDNPNSISVSGCELEAFKQ 239
QY 236 VFGSVSSRDAQIKFLNTYQNPGEKLSAYVIRLEPLQLQKVEKGAIDKDNVNQARLEQVI 295

```

```

QY 90 VLEKPTTS--DAEFLERLHLFLAREGWTQDVVARVLGFONFTPTTPGPEMPAEMLNLYLDN 147
Db 185 LLLDEPTNHLDAESVWAWLERFLHDYEGTVAITHD-----RYFLDN 225
QY 148 VTQPLVESIYWKRLTLFSGKHPRAWGNFDPWLEHTNVLBEWQSDVEKRRRLMESUR 207
Db 226 VAGWILE-----LDRGEGIP--WEGNYSWLEQEKRELEQEQATENARQKSTAKEL- 274
QY 208 GPAADVIRILKSNPAITTAECCLKALEQVFGSVESRDAQIKFLNTYQNPGEKLSAYVIR 267
Db 275 -----EWVRPNKGRQAKSMARWDELNSGYQKRNETNELFI-----PPGRLGDKVIE 326
QY 268 LEPQLQKVKEGAIDKDNVQARLEQV-IAGANHSG 302
Db 327 VQNLTKSYGDRTLDDLSPSIKGAIVGIIGANGAG 362

RESULT 11
ID KINH SYNRA STANDARD; PRT; 935 AA.
AC 043093;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Kinesin heavy chain (Synkin).
OS Syncephalastrum racemosum.
OC Eukaryota; Fungi; Zygomycota; Zygomycetes; Mucorales;
OC Syncephalastraceae; Syncephalastrum.
OX NCBI_TaxID=13706;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98273660; PubMed=9613604;
RA Grunnt M., Pistor S., Lottspeich F., Schliwa M.;
RT "Cloning and functional expression of a 'fast' fungal kinesin.";
RL FEBS Lett. 427:179-84(1998).
CC -!- FUNCTION: KINESIN IS A MICROTUBULE-ASSOCIATED FORCE-PRODUCING
CC PROTEIN THAT MAY PLAY A ROLE IN ORGANELLE TRANSPORT. ITS MOTOR
CC ACTIVITY IS DIRECTED TOWARD THE MICROTUBULE'S PLUS END. THE SPEED
CC OF THIS MOTOR IS 4-5 TIMES FASTER THAN ITS ANIMAL COUNTERPARTS.
CC -!- DOMAIN: Composed of three structural domains: a large globular N-
CC terminal domain which is responsible for the motor activity of
CC kinesin (it hydrolyzes ATP and binds microtubule), a central
CC alpha-helical coiled coil domain that mediates the heavy chain
CC dimerization; and a small globular C-terminal domain which
CC interacts with other proteins (such as the kinesin light chains),
CC vesicles and membranous organelles.
CC -!- SIMILARITY: Belongs to the kinesin-like protein family. Kinesin
CC subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; AJ225894; CAA12647.1; -
CC F1R; T51930; T51930.
CC H5SP; P33176; I8G2.
CC InterPro; IPR007420; DUF465.
CC InterPro; IPR001752; kinesin_motor.
CC Pfam; PF04325; DUF465; 1.
CC Pfam; PF00225; kinesin; 1.
CC PRINTS; PR00380; KINESINHEAVY.
CC SMART; SM00129; KUSC; 1.
CC PROSITE; PS00411; KINESIN MOTOR DOMAIN; 1.
CC PROSITE; PS00067; KINESIN MOTOR DOMAIN2; 1.
CC Motor protein; Microtubule; ATP-binding; Coiled coil.
CC DOMAIN 1 341 KINESIN-MOTOR (BY SIMILARITY).
CC FT DOMAIN 342 887 COILED COIL (POTENTIAL).
CC FT DOMAIN 888 934 GLOBULAR.
CC FT NP_BIND 87 94 ATP (POTENTIAL).

```

```

FT NP_BIND 237 244 ATP (POTENTIAL).
SQ SEQUENCE 935 AA; 105040 MW; CFBDFD33B3827E28 CRC64;

Query Match
Best Local Similarity 6.0%; Score 103; DB 1; Length 935;
Matches 69; Conservative 56; Mismatches 121; Indels 72; Gaps 15;

QY 16 NSORTLLVWGIFVNCDEAIEETLQAAMPQVSYRMGLGRMFREENAKAALLEL----- 68
Db 297 NSRTTLIINCSPSSSNEAETLSTLFGARAKSIKNKAV-----NADLSPAELKALLKKV 351
QY 69 -TGAVDYAAIPREMPGKGWVKVLPKPTSDAEFLERLHLFLAREG-WTVQDVAVLGFQ 126
Db 352 KSEAVTYQTYIAALEGEVNVWRTGTVP-----EGKWTMDKVSKGDEA 395
QY 127 NPTPTPGPEMPAEMLNLYLDNVIQPLVESIYWKRLTLFSGKHPRAWGNFDPWLEHTNE 186
Db 396 GLPPAPGFKSP-----VSDEGRPATPV-----PTEKDEREFIKRENE 435
QY 187 VLEEWQVSDVE---KRRRLMESLR---GPAADVIRILKSNPAIT--AECCLKALEQVF 237
Db 436 LMD--QISEKETELTNREKLLESREEMGYGKEQSVTKENQOMTSELSELRLQLQKV- 492
QY 238 GSVESRDAQIKFLNTYQNPGEKLSAYVIRLEPLQKV--VEKGAIDKDNVQV--ARLEQ 293
Db 493 -SYESKENAIT--VDSLKEANQDLMAELBELKNLSEMRQAHKDATDSDEKREKAKFMAQ 549
QY 294 VIAGANHSGAI---RRQL 308
Db 550 MMSGFDPGSLINDKERQI 567

RESULT 12
SYEA AERPE STANDARD; PRT; 488 AA.
ID -SYEA AERPE STANDARD; PRT; 488 AA.
AC Q9Y9I6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Phenylalanine-tRNA synthetase alpha chain (EC 6.1.1.20)
DE (Phenylalanine-tRNA ligase alpha chain) (PheRS).
DE PHES OR APE2302.
GN Aeropyrum pernix.
OS Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;
OC Desulfurococceae; Aeropyrum.
OX NCBI_TaxID=56636;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K1;
RX MEDLINE=99310339; PubMed=10382966;
RA Kawarabayasi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankai A., Kosugi H.,
RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA Yamazaki J., Kishida N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT "Complete genome sequence of an aerobic hyper-thermophilic
RT crenarchaeon, Aeropyrum pernix K1.";
RL DNA Res. 6:83-101(1999).
CC -!- CATALYTIC ACTIVITY: ATP + L-phenylalanine + tRNA(Phe) = AMP +
CC diphosphate + L-phenylalanyl-tRNA(Phe).
CC -!- SUBUNIT: Tetramer of two alpha and two beta chains (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.
CC Phe-tRNA synthetase alpha chain subfamily 2.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@sib-sib.ch).

```



```
DR InterPro; IPR000782; Bigh3_fasciclin.
DR SMART; SM00554; FAS1; 2.
DR PROSITE; PS00213; FAS1; 2.
KW Membrane; Repeat; Signal; Glycoprotein; Lipoprotein; GPI-anchor.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 404 FASCICLIN-LIKE ARABINOGLACTAN PROTEIN 1.
FT PROPEP 405 424 REMOVED IN MATURE FORM (POTENTIAL).
FT DOMAIN 25 170 FAS1 1.
FT DOMAIN 184 323 FAS1 2.
FT LIPID 404 404 GPI-anchor amidated glycine (potential).
FT CARBOHYD 26 26 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 128 128 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 160 160 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 186 186 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 240 240 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 424 AA; 44849 MW; 6B256D3CE592881B CRC64;

Query Match
Best Local Similarity 5.7%; Score 98.5; DB 1; Length 424;
Matches 79; Conservative 49; Mismatches 127; Indels 109; Gaps 18;

QY 1 MAMTLLDWCRCMDYNSORTLWGIPIVNCDEAIEE-----TLQAMPQVSYRMLGRM 54
Db 44 LQTHLAD-----EINRRRTITVCAY---DNAASALTSKGYTLSTLKNILSLHLVDY 94
QY 55 FW-----REENAKAA-LLELTGAVDYAAIPREMPGKGV-----WKVLEKPTSD 98
Db 95 FGTKKLHQIROGSAALVFOATGAA-----PGTSGFVNIITDLRGKVGFGDGDG 145
QY 99 AEFLERLHLFLARE-----GWTQDVAVRGFON-PTTPGPEMPAEMNLVLDNVIQ 151
Db 146 -----LSSFFVKSIIEVPYNIISIIQISVLPSETAAPTAA-PAEM---NLTGIMS 193
QY 152 LVESIWYKRLTLFSKGHPRAWGN-----FDPWLEHTNEVLEWQSVQVEKRRLME 204
Db 194 HGCKVFAB--TLITNPGASKTYQESLEGWTVFCPGDDAMKGFGLPKYKNLTAPKKEAFLD 251
QY 205 SLRGPAAADVIRILKSNPAITTAECLEKALEQVFGSVESRDAQIKFLNTYQNPGEKLSAY 264
Db 252 FLAVPTYYSMALKNNGPMNTL-----ATDGANKFELTVQNDGKVTIL- 295
QY 265 VIRLEPLQKVEKGAIDKNNVQARL-----EQVIA-GANHSAGAIRQLMLTGAGRGP 317
Db 296 -----KTRINTVKIVDTLIDEQPLAIYATDKVLLPKELFKASAVEAP 337
QY 318 GPKP 321
Db 338 APAP 341

RESULT 15
YM26 MYCTU
ID YM26 MYCTU STANDARD; PRT; 513 AA.
AC Q10510;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical protein Rv2226/MT2285.
GN Rv2226 OR MT2285 OR MTCY427.07.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
CX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RA MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekai F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
```

```
RA Rutter S., Seager K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the Biology of Mycobacterium tuberculosis from the
RL complete genome sequence.";
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=2206494; PubMed=12218036;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,
RA Delcher A., Uterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.W.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL J. Bacteriol. 184:5479-5490(2002).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
EMBL; 270692; CA94649.1; -.
DR EMBL; AB007073; AAK46570.1; -.
DR TIGR; F70776; P70776.
DR TIGR; MT2285; -.
DR Tuberculist; Rv2226; -.
DR InterPro; IPR008172; Adenylate_cyc.
DR InterPro; IPR007899; CHAD.
DR Pfam; PF05235; CHAD; 1.
DR Pfam; PF01928; CVTH; 1.
KW Hypothetical protein; Complete proteome.
FT CONFLICT 299 299 D -> N (IN REF. 2).
SQ SEQUENCE 513 AA; 56333 MW; 8E74BDA8945D7B99 CRC64;

Query Match
Best Local Similarity 5.7%; Score 98; DB 1; Length 513;
Matches 58; Conservative 30; Mismatches 84; Indels 82; Gaps 13;

QY 90 VLFKPTSDAEFLERLHLFLARE-----GWTQDVA---RVLGQNPPTTPGPEMPAEM 140
Db 49 VYFDTFSDH---LARNQITLRRRTGGADAGWHLKLPAGPKRTMRAPLSASGDAVPAEL 105
QY 141 LNYIL----DNVIQPLVESIWYKRLTLFSKGHPRAWGNFDPWLEHTNEVLEWQSV 196
Db 106 LDVLAIVRDQVPQVPAVARISTHRESQILYGAG-----EQW-----DALAEFCNDVDTWSAGAP 157
QY 197 EKRRLMESLRGPAADVIRILKSNPAITTAECLEKALEQVFGSVESRDAQIKFLNTYQ 256
Db 158 H-----AAGAAD-----NGPA-----EQW-----REWELELVTTDGT 185
QY 257 PGEKLSAVIRLEPLQKVEKGAIDKNNVQARLQEVIAGANHSAGAIRQLMLTGAGEG 316
Db 186 ADTKL-----LDRLANRLLDAGA-----APAGHSGKLARVLGATSPGEL 224
QY 317 P-GPKPLSVAGADP 329
Db 225 PNGPQP-----PADP 234

Search completed: September 21, 2004, 13:32:36
Job time : 10.969 secs
```

Blank

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 13:20:22 ; Search time 24.8686 Seconds
(without alignments)
1272.567 Million cell updates/sec

Title: US-10-037-860-4

Perfect score: 1729

Sequence: 1 MAMTLEEDWCKGMDVNSORT.....LTGAGEGPGPKPLSVAGADP 329

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_78.*

1: PIR1.*

2: PIR2.*

3: PIR3.*

4: PIR4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	114.5	6.6	2037	2 T16881	hypothetical prote
2	111	6.4	558	2 F81781	ABC transporter At
3	110.5	6.4	559	2 JC7667	type II ATP-bindin
4	108	6.2	558	2 H81205	ABC transporter, A
5	104.5	6.0	556	2 G64169	hypothetical prote
6	103	6.0	935	2 T51930	kinesin (imported)
7	101	5.8	473	2 B72457	probable phenylala
8	101	5.8	555	2 C87515	ABC transporter, A
9	98	5.7	513	2 P70776	hypothetical prote
10	98	5.7	554	2 C62743	ABC transporter At
11	97.5	5.6	759	2 H77443	topoisomerase IV,
12	96.5	5.6	564	2 B75607	hypothetical prote
13	95.5	5.5	555	2 H97710	ABC transporter At
14	94.5	5.5	555	2 D82290	ABC transporter, A
15	94.5	5.5	1003	2 C84601	hypothetical prote
16	94	5.4	726	2 E83712	ribonucleoside-dip
17	93	5.4	316	1 C69855	ABC transporter ho
18	93	5.4	659	2 AB0731	probable hydrolase
19	92.5	5.3	464	2 S34024	alpha-1,2-mannosyl
20	92.5	5.3	2431	1 MNWVSF	nonstructural poly
21	90.5	5.2	141	2 E81294	50S ribosomal prot
22	90.5	5.2	671	2 T04457	hypothetical prote
23	90.5	5.2	1027	2 H87316	hypothetical prote
24	90	5.2	513	2 T14864	probable monosacch
25	90	5.2	861	2 G81068	aconitate hydratase
26	90	5.2	1073	2 F89467	protein R09H3.1 [i
27	90	5.2	1660	2 T18561	vitellogenin vit-6
28	89.5	5.2	1140	2 F88349	protein F15D4.7 [i
29	89.5	5.2	1140	2 T20984	hypothetical prote

cytochrome P450 er
conserved hypothet
aconitate hydratase
probable monox prot
hypothetical prote
group II chaperoni
cyanodine receptor
transforming prote
chemotaxis-related
hypothetical prote
probable medium-ch
probable outer mem
probable leader pe
chemotaxis protein
probable transposo
translation elonga

30 89 5.1 406 1 S18531
31 89 5.1 555 2 A51074
32 89 5.1 861 2 A81801
33 88.5 5.1 330 2 H75353
34 88.5 5.1 364 2 B83152
35 88.5 5.1 552 2 JC5617
36 88.5 5.1 4872 2 S27272
37 88 5.1 640 1 A50733
38 88 5.1 1755 2 F82618
39 87.5 5.1 334 2 S25785
40 87.5 5.1 559 2 C72732
41 87.5 5.1 568 2 C72129
42 87.5 5.1 568 2 G86493
43 87.5 5.1 764 2 H87322
44 87.5 5.1 1011 2 E85054
45 87 5.0 461 1 EFH81

ALIGNMENTS

RESULT 1
T16881
hypothetical protein T14G12.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T16881
R:Wilcoxon, L.
submitted to the EMBL Data Library, November 1995
A:Description: The sequence of C. elegans coamid T14G12.
A:Reference number: Z18596
A:Accession: T16881
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2037 <MIL>
A:Cross-references: EMBL:U41268; NID:g1086843; PID:g1086848; PIDN:AAA82437.1; CESP:T14G12.5
C:Genetics:
A:Gene: CESP:T14G12.5
A:Introns: 715/2; 1484/2

Query Match 6.6%; Score 114.5; DB 2; Length 2037;
Best Local Similarity 22.0%; Pred. No. 1.5;
Matches 56; Conservative 41; Mismatches 114; Indels 43; Gaps 8;

Qy 36 BETLOAMPQVSVMGLRMFWR-EENAKAALLELTGAVDYAAIIPREMPGKGVNKLVPK 94
Db 308 EQLLPQAMLOAMTEM-GRLLSQLODQARREQTSFMNECREHLRPPAEGSIGQSAYSP 366
Qy 95 PTSDAEFLERLHLFLAREGWTQDVARVLGFQNP-PTPGPEMPAEMLYILDNVLPV 153
Db 367 DDEGEESQR-----GSSPPVQIPDSRSPGVINFETN----- 400
Qy 154 ESIWYKRLTLFSGKHPRAWRGNFDPMLEHTNEVLEWQVSDVEKRRRLMESLRGPAADV 213
Db 401 ----AKULPKFDGTGNFRANGFD-----TVVLDPRLPSTVKCLLRNHLVGNQQC 450
Qy 214 IRLKSNPNPAITAECLKALEQVFGSVESRDRQIKF--LNTYQNPQEKLSAYVIRLEPL 271
Db 451 IS--HDDPDLVAYQTTMDMLESVYKGDQGLERFRKLKFKHQSNEQMKDLTSLHLL 508
Qy 272 LQKWEKGAIKDKN 285
Db 509 VQRLVSTGLSATDD 522

RESULT 2
F81781
ABC transporter ATP-binding protein NMA2101 [imported] - Neisseria meningitidis (strain 2
C:Species: Neisseria meningitidis
C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001
C:Accession: F81781
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,

Nature 404, 502-506, 2000

A:Title: Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* Z2491.
A:Reference number: A81775; MUID:20222556; PMID:10761919
A:Accession: F81781
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-558 <PAR>
A:Cross-references: GB:AL162758; GB:AL157959; NID:g7380672; PIDN:CAB85317.1; PID:g738072
A:Experimental source: serogroup A, strain Z2491
C:Genetics:
A:Gene: NMA2101
C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology

Query Match 6.4%; Score 111; DB 2; Length 558;
Best Local Similarity 22.0%; Pred. No. 0.42;
Matches 51; Conservative 39; Mismatches 92; Indels 50; Gaps 8;

Qy 90 VLFKPKPTS--DAEFLERLHLFLAREGTVQDVARVLGFQNPPTPTPGPEMPAEMLNLYILDN 147
Db 187 LLLDEPTNHLDAESVLEWLSQFLVR-----PPGTVAVTHDRYFLDN 227

Qy 148 VIQPLVESIWYKRLTLTFSKGGHPRAWGNFDPWLEHTNEVLVEWQVSVKERRRLMESLR 207
Db 228 LAEWILE-----LDRGHGIPKMGNYSSWLEQKEKRLLENFAKSAARVAKMQBLE 277

Qy 208 -----GPAADVIRILKSNPAITTAECLEKALEQVFGSV-ESSRDAQIKFLNTYQNP 258
Db 278 WYRONAKGROAKSKARLARFEENSVEYQKNETQEIFIPVAERLGNIEVFVNYSKSGF 337

Qy 259 EKLSAYVIRLEPLQKV---VEKGAIKDKNVQARLEQVIAGANH--SGAIR 305
Db 338 DK-----VLIDDLFSKVPAGAVIGIIPNGAGKSTLFKMIAGKEQPDSEVK 384

RESULT 3

JC7667

type II ATP-binding cassette (ABC) transporter, Mac-1 - *Myxococcus xanthus*
C:Species: *Myxococcus xanthus*
C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 24-Aug-2001
C:Accession: JC7667

R:Kimura, Y.; Yamanishi, Y.; Tokumasu, Y.; Terasaka, H.; Yoshinobu, J.
J. Biochem. 129, 351-356, 2001
A:Title: Characterization of the mac-1 gene encoding a putative ABC transporter from *Myxococcus xanthus*
A:Reference number: JC7667; MUID:21125588; PMID:11226873
A:Accession: JC7667
A:Molecule type: DNA
A:Residues: 1-559 <KIM>
A:Cross-references: DDBJ:AB041227
A:Experimental source: ATCC 31046
C:Comment: This protein contributes to the export or import of a developing signal or a
C:Genetics:
A:Gene: mac-1
C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology
C:Keywords: antibiotic; sporulation

Query Match 6.4%; Score 110.5; DB 2; Length 559;
Best Local Similarity 21.5%; Pred. No. 0.47;
Matches 71; Conservative 47; Mismatches 113; Indels 99; Gaps 17;

Qy 60 NAKAALELTGAVDYAAIPREMPCKGGVWVLFKPKPTSDAEFLERLHLFLARE----- 112
Db 89 NVELGKEIKPATIDRF--NEVSAK-----FAEPMSDAE-MEKL---LAEQGRLODAI 134

Qy 113 -----GW----TVQ-----DVARVLG-----FQNP-----TPTPGPE 135
Db 135 DAVNGWELDRTIEMAMDALRLPPGDADVTKLSGGKRRVALCRILLEKPDLLLLDEPTNH 194

Qy 136 MPAEML-----NYLDNVIQPLVESIWYKRLTLFSGKGHPRAWRG 175
Db 195 LDAESVAWLQALKEYKGTIVCIITHDRYFLDNAAEWILE-----LDRGEGVP--WKG 244

Qy 176 NFQDFLEHTNEVLVEWQVSVKERRRLMESLRGPAADVIRILKSNPAITTAECLEKALEQ 235

A;Note: best homolog was a hypothetical protein from *Escherichia coli*
 C;Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology
 E;Keywords: ATP; nucleotide binding; P-loop
 F;23-236/Domain: ATP-binding cassette homology <ABC1>
 F;40-47/Region: nucleotide-binding motif A (P-loop)
 F;340-518/Domain: ATP-binding cassette homology <ABC2>
 F;357-364/Region: nucleotide-binding motif A (P-loop)

	Query Match	6.0%;	Score 104.5;	DB 2;	Length 556;
	Best Local Similarity	23.6%;	Pred. No. 1.4;		
	Matches 51; Conservative 34;	Mismatches 90;	Indels 41;	Gaps 7;	
Qy	90 VLFKPPTS--DAEFLERLHLFLAREGWTVQQVARVLGFQNPTTTPGPEMPAEMLVILDN 147	:	: :	:	:
Dd	185 LLLDEPINFHLDAESVAMLERLHDYEGVTVAITHD-----RYFDLN 225	:	:	:	:
Qy	148 VIQPLVESIWYKRLTLFSGKHPRAWRGNFPDWLEHTNTEVLBEWQSDVKERRRMLESIR 207	:	:	:	:
Dd	226 VAGWILE-----LDREGEP--WECNYGSWLQEKEKLREQEATENARQKSIAKEL- 274	:	:	:	:
Qy	208 GPAADVIRLKSNPNPATTTABCLKALEQVFQSVSSRRDAQIKELNTYNQPGEKLSAYVIR 267	:	:	:	:
Dd	275 ----BWTRQPKFGQAQSKARMARFDELNSGEYQRNETNELFI----PPGRPLUGDKVII 326	:	:	:	:
Qy	268 LEPLLQKVBSKGAIKDKNVNQARLEQV-IAGANHSG 302	:	:	:	:
Dd	327 VQNLTKSYGDRTLLDDLSFSIPKGAIVGIIGANGAG 362	:	:	:	:

```

RESULT 6
T51930
kinesin [imported] - pin mould (Syncephalastrum racemosum)
C:Species: Syncephalastrum racemosum
C>Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 03-Nov-2000
C/Accession: T51930
R:Grummt, M.; Pistor, S.; Lottspeich, F.; Schliwa, M.
FEBS Lett. 427, 79-84, 1998
A:Title: Cloning and functional expression of a 'fast' fungal kinesin.
A:Reference number: Z5870
A/Accession: T51930
A/Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-935 <GRU>
A/Cross-references: EMBL:AJ225894; PIDN:CAAL2647.1
C:Superfamily: kinesin heavy chain; kinesin motor domain homology

```

Query Match	6.0%;	Score 103;	DB 2;	Length 935;
Best Local Similarity	21.7%;	Pred. No. 4.1;		
Matches	69;	Conservative 56;	Mismatches 121;	Indels 72; Gaps 15;
QY	16	NSORTLLVWGIVNCDEAEITEETLQAAAMPQVSYRMLGEMFWRENAXAALLEL-----	68	
Db	297	NSRTILLNCSPSSYNEAEITSLTRFGARAKSINKAKV-----NADLSPAELKALLKKV	353	
QY	69	-TGAVDYAALPREMPGKGWYKLFKPKPTSDAEFLERLHLFLAREG-WTQDVAVRLGFQ	126	
Db	352	KSEAVTYQTVIAALEGEVNVWRTGGTVP-----EGKWVMDKVKSGDFA	395	
QY	127	NPTPTGCPMPDPAEMLVNIDNVTOPLVESIWYKRLTLFSGKHGPRAWRGFDPWLEHTNE	186	
Db	396	GLPPAPFGFKSP-----VSDEGSRPATPV-----PTLEKDEREEFIKRENE	435	
QY	187	VLEEMQVSDVE-----KERRLMESLR-----GPAADVIRILKSNPAITT--AECLKALEQV	237	
Db	436	LMD--QISEKETELTNREKLLESREEMGYKQEQSVTKENQOMTSELRLQLOKV--	492	
QY	238	GSVESSRDAQIKFLNTYQNFGKLSAVVIRLEPLLOKV--VEGAIDKDNVQ--ARLEQ	293	
Db	493	-SYESKENAIT--VDSLKEANQDLMAELEELKKNLSEMRQAHKQADTSDSEKKRAEKMAQ	549	
QY	294	VIAGANHSGAI-----RRQL	308	
Db	550	MMSGFDPSPGILNDKEROI	567	

RESULT 7
B72457

Probable phenylalanyl-tRNA synthetase alpha chain APE2302 - Aeropyrum pernix (strain K1)
 C:Species: Aeropyrum pernix
 C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000
 C:Accession: F72457
 R:Kawarayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takahawa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Ki
 DNA Res. 6, 83-101, 1999
 A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
 A:Reference number: A72450; MUID:99310339; PMID:10382966
 A:Accession: B72457
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-473 <KAW>
 A:Cross-references: DDBJ:AP000064; NID:g5105945; PIDN:BAA81314.1; PID:g5106003
 A:Experimental source: strain K1
 C:Genetics:
 A:Gene: APE2302
 C:Superfamily: yeast cytosolic phenylalanine-tRNA ligase beta chain

Query Match	5.8%;	Score 101;	DB 2;	Length 473;
Best Local Similarity	24.5%;	Pred. No. 2.2;		
Matches	50;	Conservative 26;	Mismatches 58;	Indels 70; Gaps 12;
QY	31	DEAE-----IETLQAAAPQVSRYM-----LGRMFWRERNAKAALEL-----	68	
Db	116	DQAEALKTITPLKLLLENVAGSGKPTVGDELLREALSGLIRRRARISIVLRKVNPAEA	175	
QY	69	--TGAVDYAAIPREMFGKGWVKVLPKP-----PTSD--AEFLERLHLFLAR	111	
Db	176	LAPARVEAAVLITDML-KSGEWRLRPKYNVKAEPVLPARRHFLAEFTLERL-----	228	
QY	112	EGMTVOODVARVLGFQNPPTPFGPEMPAEMLYIL-----DNVIQPLVESIWVK--RLTLF	164	
Db	229	-----RDILRELGFRE-----VRGPLELELEFNFDVLFQADHPAREIHSDLMKSPRRGDL	280	
QY	165	SGKG-----HPRAWGNFDP	179	
Db	281	SGYSDLIVERVASVHERGWKYSRP	304	

RESULT 8

C87515
 ABC transporter, ATP-binding protein CC2148 [imported] - Caulobacter crescentus
 C:Species: Caulobacter crescentus
 C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 22-Oct-2001
 C:Accession: C87515
 R:Rierman, W.C.; Feldblum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
 B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon-
 N, J.; Emolova, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
 Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
 A:Title: Complete Genome Sequence of *Caulobacter crescentus*.
 A:Reference number: AB7249; MOID:21173698; PMID:11259647
 A:Accession: C87515
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-555 <STO>
 A:Cross-references: GB:AA005673; NID:gl3423641; PIDN:AAK24119.1; GSPDB:GNC00148
 C:Genetics:
 A:Gene: CC2148
 C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology

	Query Match	5.8%	Score 101;	DB 2;	Length 555;
Best Local Similarity	23.3%;	Pred. No. 2.8;			
Matches	78;	Conservative	47;	Mismatches 118;	Indels 92; Gaps 19;
QY	20	TLVWG-IPVNCDEABIEBTLQAAMPQVSRYMLGRMFWEENAKAALLELTGAVDAAIP	78		
DB	84	TLVWNGVIADCEKQIFDRYNALAAQGEETDELM--EENTK--LQEVIDARD----	134		

Query Match 5.6%; Score 97.5; DB 2; Length 759;
Best Local Similarity 23.1%; Pred. No. 8.6;
Matches 71; Conservative 41; Mismatches 108; Indels 87; Gaps 17;

QY 17 SQTLLVMGIPVNCDEARIEETLQAAMPQVSVMGLGRMFWEENAKAALLELTG----AV 72
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 107 SQRIPLVEG---QGNFGNIDGSAAAMRYTECKM-----TEAALLLDGDIEDAV 153
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 73 DY-----AAIREMP-----GKGG--VMKVLFKPPTSDAEFLERLHLFLAREGW 114
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

DB 154 DFRPTYDQDBEPVVLPSPGFNFLLANGSSGIAGVMATSIIPHNAELIDACQLLANPDA 213
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 115 TVQDVARVLGQNTPPTPGPEMPAEMLNIYLDNVIOPLVESIWTYKRITLFSKGKHPR--- 171
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

DB 214 TTADILE-----KVGPDFFPGGV-----IVESRASLIETTYTGRGGVMRA 255
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 172 AW-----RGNFD-----PWLEHTNEVLEEWOVSDV--EKRRIMESLRGAADVIRIL 217
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

DB 256 KWEXEDTGRGYQIVVTETPIQOVKKSDLVE--QLADLDSKAALIGVDREDSREDRLV 313
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 218 ---KSNN--PAITTAECCLKALEQVFSGVESRDQAIKPLNTYONPG-----EKL SAYVIR 267
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

DB 314 LEPKSKNVEPEVLMESELFK-----LSALESFPVNINVLDRGTGPMVGIIKQALMAFLAH 368
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 268 LEPLOK 274
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

DB 369 RREVLR 375
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 12
H75607
hypothetical protein - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
C:Accession: H75607
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.C.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma,
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A;Reference number: A75250; UID:20036896; PMID:10567266
A;Accession: H75607
A>Status: preliminary
A:Molecule type: DNA
A;Residues: 1-564 <WHI>
A;Cross-references: GB:AE001862; GB:AE001825; NID:g6460468; PIDN:AAF12319.1; PID:g646061
A;Experimental source: strain R1
C:Genetics:
A;Gene: DRA0125
A;Map position: 2

[illegible]

C;Genetics:
A;Gene: VC0698
A;Map position: 1
C;Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology

Query Match 5.5%; Score 94.5; DB 2; Length 555;
Best Local Similarity 21.5%; Pred. No. 9.6;
Matches 73; Conservative 54; Mismatches 116; Indels 97; Gaps 16;

QY 31 DEAE-IETLQAAPOVSY--RMLGRMF-----WRENAKAALE-----L 68
DB 81 DESKTVREVEAADVAYALKKELDEVAYAYABPDADFALAKEQGELEALIQAGHNL 140
QY 69 TGAVDYAAIPRMWG-----KGGWK-----VLFKPPTS--DAEFLER 104
DB 141 DNALERAADALRLPEWDAQVYLSGERRRVAICRLLENPDMLLDEPTNHLDAESVAV 200
QY 105 LHLFLAREGWTVQDVARVLGFQNPPTTPGPMPAEMLYILDNVIQPLVESIWYKRLTLF 164
DB 201 LERFLVDYAGTVVAITHD-----RYFLDNAAGWILE-----LD 233
QY 165 SKKGHPRAWGNFDPMLHEHTNEVLEBWQSVDEKRRLEMSLRGPAADVIRILKSNPAI 224
DB 234 RGEGIP--WQNYTSLMLQKQDARLQ--EASQEKARQKTIK-----ELEWVR---QNPKG 282
QY 225 TTAECCLKALEQFGSVSSRDAQIKFLN--TYQNPGEKLSAYVIRLEPLLQKWEKGAIDK 283
DB 283 RQAKS--KARMAFEELONTEHOKENEINELFIPPERLGDKVIEKNLTKSFQDGRVLDD 341
QY 284 DNVNQARLEQVITAGANHSAGIRQLWLITGAGEGPGPPLS 323
DB 342 LSFSPMK-----GAIVGIIGNGAGKSTLFKMLS 370

RESULT 15
C84601
hypothetical protein At2g21440 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C;Accession: C84601
R.Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.B.; Umayam, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: C84601
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1003 <STO>
A;Cross-References: GB:AE002093; NID:g4567275; PIDN:AAD23688.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g21440
A;Map position: 2

Query Match 5.5%; Score 94.5; DB 2; Length 1003;
Best Local Similarity 19.3%; Pred. No. 23;
Matches 59; Conservative 42; Mismatches 74; Indels 131; Gaps 12;

QY 18 QRTLLVWGIPVNCDEABIE-----ETLQAAAPQVSYRMLGRMFWRRENAKAALEL 68
DB 560 ERTLFTNLFPDVTKEVKQRPFTVFGVESLSVLHKVTKRPEGTAFAVKFKTADASVAAL 619
QY 69 TGAVDYAAIPREMPKG-GVWNVLFKPTSDAEF-----LERLHLFLAREGWTVQDVA 120
DB 620 SAADTASGVGLLKGRQLNVNRAVQKKAADIELKKTTEKNVDHRLNYLAKEG----- 672
QY 121 RVLGFQNPPTTPGPMPAEMLYILDNVIQPLVESIWYKRLTLFSGKGHPRAWGNFDPW 180
DB 673 -----QILDDT--FAAEGV----- 684
QY 181 LEHTNEVLEBWQSVDEKRRRL-----MESLRGPAADVIR-----ILKSNPAITTAEC 229

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 13:35:22 ; Search time 76.0979 Seconds
(without alignments)
1194.261 Million cell updates/sec

Title: US-10-037-860-11

Perfect score: 1462

Sequence: 1 VQKGGVWKVFKTPNQDTE.....SIEPPEERDGYRNWGGD 283

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1342398 seqs, 321133274 residues

Total number of hits satisfying chosen parameters: 1342398

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published Applications AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	DB ID	Description
1	1462	100.0	283	13	US-10-037-860-11
2	755	51.6	149	13	US-10-037-860-9
3	628	43.0	353	9	US-09-965-529-7
4	628	43.0	353	10	US-09-969-680A-7
5	618.5	42.3	463	13	US-10-037-860-13
6	597	40.8	452	16	US-10-408-765A-2385
7	596.5	40.8	351	9	US-09-965-529-1
8	596.5	40.8	351	10	US-09-969-680A-1
9	596.5	40.8	351	12	US-09-804-014A-16
10	596.5	40.8	351	15	US-10-341-434-10
11	593	40.6	195	13	US-10-037-860-7
12	564	38.6	329	13	US-10-037-860-4
13	562.5	38.5	318	12	US-09-804-014A-40
14	560	38.3	321	12	US-09-804-014A-39
15	550	37.6	312	12	US-09-804-014A-73

16	550	37.6	312	12	US-09-804-014A-74	Sequence 74, Appl
17	475.5	32.5	399	15	US-10-094-749-1978	Sequence 1978, Ap
18	394	26.9	403	15	US-10-094-466-38	Sequence 38, Appl
19	378.5	25.9	337	12	US-10-296-115-1208	Sequence 1208, Ap
20	215	14.7	120	12	US-09-804-014A-42	Sequence 42, Appl
21	156.5	10.7	204	14	US-10-029-386-33747	Sequence 33747, A
22	123	8.4	120	12	US-09-804-014A-41	Sequence 41, Appl
23	117	8.0	538	16	US-10-408-765A-2992	Sequence 2992, Ap
24	113	7.7	2383	14	US-10-082-830-260	Sequence 260, App
25	110.5	7.6	584	15	US-10-221-278-355	Sequence 355, App
26	110.5	7.6	584	15	US-10-291-172-355	Sequence 355, App
27	107	7.3	1031	11	US-09-764-875-686	Sequence 686, App
28	107	7.3	1035	15	US-10-158-057-197	Sequence 137, App
29	107	7.3	1459	16	US-10-408-765A-2246	Sequence 2246, Ap
30	106	7.3	542	12	US-10-205-331-57	Sequence 57, Appl
31	105.5	7.2	758	12	US-10-282-122A-67949	Sequence 67949, A
32	102.5	7.0	879	12	US-10-282-122A-60655	Sequence 60655, A
33	102	7.0	116	9	US-09-864-761-34645	Sequence 34645, A
34	102	7.0	225	10	US-09-764-891-4172	Sequence 4172, Ap
35	101.5	6.9	788	14	US-10-128-714-8204	Sequence 8204, Ap
36	100	6.8	750	14	US-10-410-681-12	Sequence 12, Appl
37	100	6.8	1082	16	US-10-437-963-162190	Sequence 162190,
38	99.5	6.8	860	12	US-10-072-012-838	Sequence 838, App
39	99.5	6.8	860	12	US-10-037-417-59	Sequence 59, Appl
40	99.5	6.8	860	15	US-10-080-334-166	Sequence 166, App
41	99	6.8	342	12	US-10-425-114-71718	Sequence 71718, A
42	99	6.8	750	12	US-10-424-599-268662	Sequence 268662,
43	98.5	6.7	520	15	US-10-220-381-12	Sequence 12, Appl
44	98	6.7	1070	14	US-10-420-845-22	Sequence 22, Appl
45	98	6.7	1504	9	US-09-932-145-7	Sequence 7, Appl

ALIGNMENTS

RESULT 1
US-10-037-860-11
; Sequence 11, Application US/10037860
; Publication No. US20020123114A1
; GENERAL INFORMATION:
; APPLICANT: Jerome B. Posner
; APPLICANT: Josep O. Dalmau
; APPLICANT: Myrna R. Rosenfeld
; TITLE OF INVENTION: MA FAMILY POLYPEPTIDES AND ANTI-MA
; TITLE OF INVENTION: ANTIBODIES
; FILE REFERENCE: 2581.1004-004
; CURRENT APPLICATION NUMBER: US/10/037,860
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 09/189,527
; PRIOR FILING DATE: 1998-11-10
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 283
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-037-860-11

Query Match	100.0%;	Score 1462;	DB 13;	Length 283;
Best Local Similarity	100.0%;	Pred. No. 9.6e-128;		
Matches 283;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	VQKGGVWKVFKTPNQDTEFLERLNLFLKEGQTVSGMFRALQGEVSPATVPCISPEL	60	
Db	1	VQKGGVWKVFKTPNQDTEFLERLNLFLKEGQTVSGMFRALQGEVSPATVPCISPEL	60	
QY	61	LAHLGQAHAPQPLPMRYKLRVFGSAVPAPPEESFEVWLEQATEIVKWPVTEAE	120	
Db	61	LAHLGQAHAPQPLPMRYKLRVFGSAVPAPPEESFEVWLEQATEIVKWPVTEAE	120	
QY	121	KRWLAESLRGPDLMHIVQADNPISVIECELEAFQVFGSLESRRTAQVRYLKTQEE	180	
Db	121	KRWLAESLRGPDLMHIVQADNPISVIECELEAFQVFGSLESRRTAQVRYLKTQEE	180	

Qy	1.81	GEKSYAVYRLRLETTLLRKAVEKRAIPRRITADQVRLEQVNMAGATLNQMLMCRLELKDQGGP	240
Db	1.81	GEKSYAVYRLRLETTLLRKAVEKRAIPRRITADQVRLEQVNMAGATLNQMLMCRLELKDQGGP	240
Qy	241	PSFLELMKVIRPEEEEEASFPNESIEEPPEERDGYGRWNHGGD	283
Db	241	PSFLELMKVIRPEEEEEASFPNESIEEPPEERDGYGRWNHGGD	283

RESULT 2
US-10-037-860-9
Sequence 9, Application US/10037860
Publication No. US20020123114A1
GENERAL INFORMATION:
APPLICANT: Jerome B. Posner
APPLICANT: Josep O. Dalmau
APPLICANT: Myrna R. Rosenfeld
TITLE OF INVENTION: Ma FAMILY POLYPEPTIDES AND ANTI-Ma
TITLE OF INVENTION: ANTIBODIES
FILE REFERENCE: 2581.1004-004
CURRENT APPLICATION NUMBER: US/10/037, 860
CURRENT FILING DATE: 2001-01-04
PRIOR APPLICATION NUMBER: 09/189, 527
PRIOR FILING DATE: 1998-11-10
NUMBER OF SEQ ID NOS: 14
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 9
LENGTH: 149
TYPE: PRT
ORGANISM: homo sapiens
US-10-037-860-9

Query Match	51.6%;	Score 755;	DB 13;	Length 149;
Best Local Similarity	98.7%;	Prod. No. 3.2e-62;		
Matches 147;	Conservative	1;	Mismatches	1;
			Indels	0;
			Gaps	0
Qy	135	DLMHVQADNPISVEECLAEAFQVFGSLESRRTAQVRLKTYQEEGEKVSAYVLRLETL	194	
Db				
Qy	1	DLMHVQADNPISVEECLAEAFQVFGSLESRRTAQVRLKTYQEEGEKVSAYVLRLETL	60	
Db				
Qy	195	LRKAVEKRAIPRIADQVRLEQVMAGATNQLWCRLRELKQGGPPSPFLELMKVIREEE	254	
Db	61	LRRAVEKRAIPRIADQVRLEQVMAGATNQLWCRLRELKQGGPPSPFLELMKVIREEE	120	
Qy	255	EEEAASFENESTPEERDCYGRWNHEGDD	283	
Db	121	EEEAASFENESTPEERDCYGRWNHEGDD	149	

```

RESULT 3
US-09-965-529-7
; Sequence 7, Application US/09965529
; Publication No. US20020182671A1
; GENERAL INFORMATION:
; APPLICANT: LAL, Preeti
; APPLICANT: YUE, Henry
; APPLICANT: TANG, Y. Tom
; APPLICANT: BANDMAN, Olga
; APPLICANT: BURFORD, Neil
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: LU, Dyung Aina M.
; APPLICANT: PATTERSON, Chandra
; TITLE OF INVENTION: MEMBRANE ASSOCIATED PROTEINS
; FILE REFERENCE: PP-0731 USA
; CURRENT APPLICATION NUMBER: US/09/965,529
; CURRENT FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/149,641; 60/164,203; PCT/US00/22315
; PRIOR FILING DATE: 1999-08-17; 1999-11-09; 2000-08-14
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PERL Program
; SEQ ID NO 7

```



```

; PRIOR APPLICATION NUMBER: 60/189,139
; PRIOR FILING DATE: 2000-03-14
; PRIOR APPLICATION NUMBER: 60/189,140
; PRIOR FILING DATE: 2000-03-14
; PRIOR APPLICATION NUMBER: 60/190,401
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/190,231
; PRIOR FILING DATE: 2000-03-17
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 40
; LENGTH: 318
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: VARIANT
; FEATURE: (20)
; OTHER INFORMATION: Wherein Xaa is any amino acid as defined in the
; OTHER INFORMATION: specification
US-09-804-014A-40

Query Match      38.5%; Score 562.5; DB 12; Length 318;
Best Local Similarity 48.9%; Pred. No. 8.1e-44;
Matches 113; Conservative 45; Mismatches 64; Indels 9; Gaps 3;

QY      3  GGGGKWKVIFKTPNQDTEFLERLNLFLEKEGQTVSGMERALGOEGVSPATVPCISPPELLA 62
Db      83  GGGGKWKVIFKTPNDAEFLERLNLFLEKEGQTVSGMERALGOEGVSPATVPCISPPELLA 138

QY      63  HLLGQAMAHAPQLLP-PMRYRKLRFVSGSAVPAPEEPEFVWLEQATEIVKQWVPTAEK 121
Db      139  EMLNVLNDVQLPVSIVYKRLTFSGKHPRAMRGNFDPLWHTNEVLEWQVSDVEK 198

QY      122  KRWLAESLRGPDALDMLHIVQADNPSISVBECELEAFQVFGSLERRTAQVRVLYKTYQEG 181
Db      199  RRRLESLRGPALDVIRLKNPNPAITTAELKALEQVFGSVESRDQIKFLNTYQNGF 258

QY      182  EKEKVSAYVLRLETLRKAVEKRAIPRIADQVRLEQVMAGA---TLNQMLW 228
Db      259  EKL SAYVIRLEPLLQKVVRKGAIDKQNVNQARLEQVIAGNHSGAIRRQLW 309

RESULT 14
US-09-804-014A-39
; Sequence 39, Application US/09804014A
; Publication No. US20030064489A1
; GENERAL INFORMATION:
; APPLICANT: Li, Li
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Vernet, Corine
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard
; APPLICANT: Spaderna, Steven
; APPLICANT: Majumder, Kumud
; TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-721 US
; CURRENT APPLICATION NUMBER: US/09/804,014A
; CURRENT FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: 60/188,316
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 60/188,277
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 60/189,139
; PRIOR FILING DATE: 2000-03-14
; PRIOR APPLICATION NUMBER: 60/189,140
; PRIOR FILING DATE: 2000-03-14
; PRIOR APPLICATION NUMBER: 60/190,401
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/190,231
; PRIOR FILING DATE: 2000-03-17
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 39
; LENGTH: 312
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-804-014A-73

Query Match      37.6%; Score 550; DB 12; Length 312;
Best Local Similarity 50.9%; Pred. No. 1.2e-42;
Matches 113; Conservative 41; Mismatches 66; Indels 2; Gaps 2;

QY      1  VQGGKWKVIFKTPNQDTEFLERLNLFLEKEGQTVSGMERALGOEGVSPATVPCISPPELL 60
Db      82  IPGGGIWRVIFKPPDPDNTFLSRNLNEFLAGEGTMVGLSRALGHENGLDPEQGMIPDM 141
```

```

; LENGTH: 321
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-804-014A-39

Query Match      38.3%; Score 560; DB 12; Length 321;
Best Local Similarity 48.6%; Pred. No. 1.4e-43;
Matches 119; Conservative 44; Mismatches 72; Indels 10; Gaps 4;

QY      1  VQGGKWKVIFKTPNQDTEFLERLNLFLEKEGQTVSGMERALGOEGVSPATVPCISPPELL 60
Db      82  IPGGGIWRVIFKPPDPDNTFLSRNLNEFLAGEGTMVGLSRALGHENGLDPEQGMIPDM 141

QY      61  LAHLIGQAMAHAPQLLP-PMRYRKLRFVSGSAVPAPEEPEFVWLEQATEIVKQWVPTAE 119
Db      142  WAPMLAQAL-EALQPALQCLVKYKLRVFGSPPEPEEGEFGRMFHTTQMILKAWQVQDV 200

QY      120  EKKRWLAESLRGPDALDMLHIVQADNPSISVBECELEAFQVFGSLERRTAQVRVLYKTYQ 179
Db      201  EKKRRLLESRLRGPALDVIRLKNPNPLITVDECLQALEEVFGVTONPRELQVKYLTITYQK 260

QY      180  EGEKVSAYVLRLETLRKAVEKRAIPRIADQVRLEQVMAGA---TLNQMLWCLRLRLKLD 236
Db      261  DEEKL SAYVLRLEPLLQKVVRKGAIDKQNVNQARLEQVIAGVHKTIRREL-----NUP 315

QY      237  QGPPPP 241
Db      316  DGAPAP 320

RESULT 15
US-09-804-014A-73
; Sequence 73, Application US/09804014A
; Publication No. US20030064489A1
; GENERAL INFORMATION:
; APPLICANT: Li, Li
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Vernet, Corine
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard
; APPLICANT: Spaderna, Steven
; APPLICANT: Majumder, Kumud
; TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-721 US
; CURRENT APPLICATION NUMBER: US/09/804,014A
; CURRENT FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: 60/188,316
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 60/188,277
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 60/189,139
; PRIOR FILING DATE: 2000-03-14
; PRIOR APPLICATION NUMBER: 60/189,140
; PRIOR FILING DATE: 2000-03-14
; PRIOR APPLICATION NUMBER: 60/190,401
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/190,231
; PRIOR FILING DATE: 2000-03-17
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 73
; LENGTH: 312
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-804-014A-73

Query Match      37.6%; Score 550; DB 12; Length 312;
Best Local Similarity 50.9%; Pred. No. 1.2e-42;
Matches 113; Conservative 41; Mismatches 66; Indels 2; Gaps 2;

QY      1  VQGGKWKVIFKTPNQDTEFLERLNLFLEKEGQTVSGMERALGOEGVSPATVPCISPPELL 60
Db      82  IPGGGIWRVIFKPPDPDNTFLSRNLNEFLAGEGTMVGLSRALGHENGLDPEQGMIPDM 141
```


B/ank

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 13:20:22 ; Search time 21.0409 Seconds
(without alignments)
694.369 Million cell updates/sec

Title: US-10-037-860-11
Perfect score: 1462
Sequence: 1 VQKGGVWVIFKTPNQDTE.....SIEPPERDGYGRWNHGDD 283

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents AA.*
- 1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
 - 2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
 - 3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
 - 4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
 - 5: /cgn2_6/ptodata/2/iaa/PTUS_COMB.pep.*
 - 6: /cgn2_6/ptodata/2/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	618.5	42.3	462	4	US-09-189-527-13
2	593	40.6	195	4	US-09-189-527-7
3	564	38.6	329	4	US-09-189-527-4
4	100	6.8	750	4	US-09-585-173B-12
5	98	6.7	1070	3	US-08-922-635-22
6	98	6.7	1504	4	US-09-364-206-2
7	97.5	6.7	651	3	US-08-650-766-6
8	97.5	6.7	651	3	US-08-922-635-5
9	97.5	6.7	651	4	US-09-389-487-6
10	95	6.5	754	4	US-09-585-173B-51
11	94	6.4	1560	4	US-09-264-512B-2
12	93.5	6.4	331	3	US-08-556-419-25
13	93	6.4	1805	1	US-07-833-313-2
14	92	6.3	736	4	US-08-252-991A-19048
15	92	6.3	1898	1	US-08-056-200-94
16	92	6.3	1898	2	US-08-800-644-94
17	91.5	6.3	300	4	US-09-252-991A-23947
18	91	6.2	497	4	US-09-345-473E-8
19	91	6.2	518	3	US-09-329-418-3
20	91	6.2	518	3	US-09-329-418-4
21	91	6.2	518	3	US-09-329-418-5
22	91	6.2	518	3	US-09-329-418-9
23	91	6.2	518	3	US-09-531-914-3
24	91	6.2	518	3	US-09-531-914-4
25	91	6.2	518	3	US-09-531-914-5
26	91	6.2	518	3	US-09-531-914-9
27	91	6.2	1786	3	US-08-973-462-8

28	90.5	6.2	420	3	US-09-329-418-8	Sequence 8, Appli
29	90.5	6.2	420	3	US-09-531-914-8	Sequence 8, Appli
30	90	6.2	257	4	US-09-107-532A-6287	Sequence 6287, Ap
31	90	6.2	592	2	US-08-736-770-6	Sequence 6, Appli
32	90	6.2	592	4	US-09-702-705-1809	Sequence 1809, Ap
33	90	6.2	592	4	US-09-736-457-1809	Sequence 1809, Ap
34	90	6.2	592	4	US-09-643-657-4	Sequence 4, Appli
35	90	6.2	592	4	US-09-671-325-1809	Sequence 1809, Ap
36	89.5	6.1	800	4	US-09-555-790A-2	Sequence 2, Appli
37	89	6.1	344	1	US-08-843-993-3	Sequence 3, Appli
38	89	6.1	344	3	US-09-059-520A-3	Sequence 3, Appli
39	89	6.1	344	3	US-09-334-275-3	Sequence 3, Appli
40	89	6.1	1581	4	US-09-866-108A-15754	Sequence 15754, A
41	88.5	6.1	182	2	US-08-602-941-2	Sequence 2, Appli
42	88.5	6.1	182	3	US-08-961-264-2	Sequence 2, Appli
43	88.5	6.1	182	4	US-09-442-099A-2	Sequence 2, Appli
44	88.5	6.1	182	4	US-09-442-099A-17	Sequence 17, Appli
45	88.5	6.1	182	4	US-09-612-342-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-09-189-527-13
; Sequence 13, Application US/09189527A
; Patent No. 6387639
; GENERAL INFORMATION:
; APPLICANT: Jerome B. Posner
; APPLICANT: Josep O. Dalmau
; APPLICANT: Myrna R. Rosenfeld
; TITLE OF INVENTION: Ma Family Polypeptides and Anti-Ma
; TITLE OF INVENTION: Antibodies
; FILE REFERENCE: SLK98-01
; CURRENT APPLICATION NUMBER: US/09/189,527A
; CURRENT FILING DATE: 1998-11-10
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 13
; LENGTH: 462
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-189-527-13

Query Match 42.3%; Score 618.5; DB 4; Length 462;
Best Local Similarity 50.2%; Pred. No. 1.9e-56;
Matches 135; Conservative 44; Mismatches 85; Indels 5; Gaps 3;
QY 1 VQKGGVWVIFKTPNQDTEFLRLNLFLEKEGQTUSGMPALQGEVSPATVPCISPEL 60
Db 76 IPKGGPWEVIVPRNSDGEFLRLNLFLEKEERTVSDMNRVLGSDTNCAPRTVISPEF 135
QY 61 LAHLLGQMAHAPOPLI-PMRYKLRVFGSVAPEEESFEVWLEQATEIVKEMPVTEA 119
Db 136 WT--WAQTLCGAQVPLLEQMLYRELRFVSGNTTIPGALAFDAWLEHTTELMQWQPEG 193
QY 120 EKRWLAESLRGPAIDLMHIVQADNPISVVECLEAFQVFGSLERRTAQVRLKTYQE 179
Db 194 EKRRIMECLRGPAIQVSGLRASNASITVEECLAAQQVFGPVESHKIAQVVKLCKAYQE 253
QY 180 EGEKVSAYVLRLLETLARKAVEKAIPIRRADQVRLQVQVAGATLQMLWCRRLBKDDQP 239
Db 254 AGEKVSFVLRLEPLQRAVENNVSRNNVQTRLRKRVLSGATLPDKLRDKLMKQRRK 313
QY 240 PPSFLEIMKVIREEEEESAF--ENESIE 266
Db 314 PPGFLALVKLLREEEWEATLGPDRSLE 342

RESULT 2
US-09-189-527-7
; Sequence 7, Application US/09189527A
; Patent No. 6387639

```
GENERAL INFORMATION:
; APPLICANT: Jerome B. Posner
; APPLICANT: Josep O. Dalmau
; APPLICANT: Myrna R. Rosenfeld
; TITLE OF INVENTION: Ma Family Polypeptides and Anti-Ma
; TITLE OF INVENTION: Antibodies
; FILE REFERENCE: SLK98-01
; CURRENT APPLICATION NUMBER: US/09/189,527A
; CURRENT FILING DATE: 1998-11-10
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 195
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-189-527-7

Query Match      40.6%; Score 593; DB 4; Length 195;
Best Local Similarity 98.3%; Pred. No. 2,5e-54;
Matches 113; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VQGGGWKVIKTPNQDTFELRLNLFLEKGGQTVSGMFRALGQEGVSPATVPCISP 60
Db 81 VQGGGWKVIKTPNQDTFELRLNLFLEKGGQTVSGMFRALGQEGVSPATVPCISP 140
QY 61 LAHLGQAMAHAPQPLPMRYKRLRVFSGSAVPAPEESFEVWLEQATEIVKEWP 115
Db 141 LAHLGQAMAHAPQPLPMRYKRLRVFSGSAVPAPEESFEVWLEQATEIVKEWP 195

RESULT 3
US-09-189-527-4
; Sequence 4, Application US/09189527A
; Patent No. 6387639
; GENERAL INFORMATION:
; APPLICANT: Jerome B. Posner
; APPLICANT: Josep O. Dalmau
; APPLICANT: Myrna R. Rosenfeld
; TITLE OF INVENTION: Ma Family Polypeptides and Anti-Ma
; TITLE OF INVENTION: Antibodies
; FILE REFERENCE: SLK98-01
; CURRENT APPLICATION NUMBER: US/09/189,527A
; CURRENT FILING DATE: 1998-11-10
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 329
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-189-527-4

Query Match      38.8%; Score 564; DB 4; Length 329;
Best Local Similarity 47.0%; Pred. No. 6.2e-51;
Matches 117; Conservative 47; Mismatches 73; Indels 12; Gaps 4;

QY 3 GKGGWVKVIFKTPNQDTFELRLNLFLEKGGQTVSGMFRALGQEGVSPATVPCISP 62
Db 83 GKGGWVKVIFKTPNSDAEFLERLHLFLAREGTVQDVARVLGQNPTTP-----GPEMPA 138
QY 63 HLHGQAMAHAPQPL-PMRYKRLRVFSGSAVPAPEESFEVWLEQATEIVKEWPVTEAK 121
Db 139 EMLNYILDNIQPLVESIWYKRLTFSGKHPRAWGNFDPWLEHTNEVLEEWQVSDVEK 198
QY 122 KRMIAESLRGPDALDMLHIVQADNPSISVECLEAFKQVFGSLERRTAQVRLKTYOEG 181
Db 199 RRRLMESLRGPAADVIRILKSNPAITTAELKALQEVFGSVSSRDAQKFLNTYONGP 258
QY 182 EKVSAYVRLTLLRKAVEKRAIPRIADQVRLQVFNWAGA-----TLNQMLWCLRLKQ 237
Db 259 EKLSAYVIRLEPLLOKVVEKGAIDKDNVNQARLEQVIAGANHSARIRQLWL---TCAGE 315
QY 238 GPPPSFLEL 246
| | | | |
```

```
Db 316 GPGPKPLSV 324

RESULT 4
US-09-585-173B-12
; Sequence 12, Application US/09585173B
; Patent No. 6570063
; GENERAL INFORMATION:
; APPLICANT: Butler, Karlene
; APPLICANT: Famodu, Onolayo O.
; APPLICANT: Gutteridge, Steven
; APPLICANT: Maxwell, Carl
; TITLE OF INVENTION: Magnesium Chelataase
; FILE REFERENCE: BB1370 US NA
; CURRENT APPLICATION NUMBER: US/09/585,173B
; CURRENT FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: US 60/137,461
; PRIOR FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 12
; LENGTH: 750
; TYPE: PRT
; ORGANISM: Glycine max
US-09-585-173B-12

Query Match      6.8%; Score 100; DB 4; Length 750;
Best Local Similarity 23.0%; Pred. No. 0.14;
Matches 64; Conservative 40; Mismatches 94; Indels 80; Gaps 13;

QY 21 FLERLNLFEKEG-----QTVSGMFRALGQEGVS-----PATVPCISPE---LLAHL 65
Db 214 YVDEINLL--DEGISNLLNLVSEGVNTVEREGISFKHPCREPLLIATYNPEEGAVREHLL 271
QY 66 GQAMAHAPQPLPMRYKRLRVFSGSAVPAPEES--FEVWLEQ-----ATEIVKE 113
Db 272 DRIAINLSAD-LPMSFENRVAAVGATATEQENSSQVFVEETDNAKTQIILAREYKLD 330
QY 114 WPVTEAEKKRWLAESLRGPDALDMLHIVQADNPSISVECLEAFKQVFGSLERRTAQVRY 173
Db 331 VTINRDQKLVIVLEALRGCGQG--H--RAELFAARVAKCLAA-----LEGRE----- 373
QY 174 LKTYOEGEKVSAVYVRLTLLRKAVEKRAIPRIADQVRLQVFNWAGATINQMLWCLRE 233
Db 374 -KYYVDD-----LKKAVELVILPSIITESPPDQ----- 401
QY 234 LKXQGGPPPSFLELMKVIREEEESAEFENESIEPEER 271
Db 402 -QNQPPPPPPPPONQSGEGEONEEEDDDKDENEEQ 438

RESULT 5
US-08-922-635-22
; Sequence 22, Application US/08922635A
; Patent No. 6033871
; GENERAL INFORMATION:
; APPLICANT: PILETZ, John E.
; APPLICANT: IVANOV, Tina R.
; TITLE OF INVENTION: DNA MOLECULES ENCODING IMIDALINE RECEPTIVE POLYPEPTIDES
; TITLE OF INVENTION: AND POLYPEPTIDES ENCODED THEREBY
; FILE REFERENCE: Corrected Sequence Listing
; Patent No. 6033871
; CURRENT APPLICATION NUMBER: US/08/922,635A
; CURRENT FILING DATE: 1997-09-03
; EARLIER APPLICATION NUMBER: 08/650,766
; EARLIER FILING DATE: 1996-05-20
; EARLIER APPLICATION NUMBER: 60/012,600
; EARLIER FILING DATE: 1996-03-01
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
; LENGTH: 1070
; TYPE: PRT
```

```

; ORGANISM: Homo sapiens
US-08-922-635-22

Query Match      6.7%; Score 98; DB 3; Length 1070;
Best Local Similarity 22.6%; Pred. No. 0.39;
Matches 65; Conservative 33; Mismatches 103; Indels 86; Gaps 14;

QY 30 EKEGQTVSGMFRAL-----GQGVSPATVPCISPELLAHLGQAMAHAP 73
   ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
Db 10 EKELDTVE-VLKATQKAKEVSKLSNPKKGGESRLSAAFCIRPSSPPTVAPASALP 68
   ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
QY 74 QPLLPMYRKLRFVSGSNAVPAPEESFEVWLEQATEIVKE-WPVTEAEKKRWLAESLRG- 131
   ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
Db 69 QPIL-----SNQIMFVQEEALASSLSDSLTPEHQPIAQG-----CSDSLES1 113
   ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
QY 132 PA-----LDLHVIQADNPSISVECLFAFKQVF---CSLSRRRTAQVRYLKYQEE 180
   ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
Db 114 PAGQAASDDLDRDVPAGVGAGSP-----EHAPEVQVVPVPGSGQIIFLPFTCIGYTATNQD- 167
   ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
QY 181 GEKVSAYVLRLETLRLKAVEKRAIPRIADQVRLEQVMAGATLNQMLWCLRELKDKQGP 240
   ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
Db 168 -----FIQLSTLIRQAIE-RQLP-----AWIEAANQREBQGG 199
   ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
QY 241 PSFLELMKVIREEEERASPFENESIE-----EPEERDGYGRWNHGGDD 283
   ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
Db 200 EQGEE---EDEEEEDVAENRYFEMGPPDVEEEGGGQGEEREE 243
   ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :

```

```

RESULT 6
US-09-364-206-2
; Sequence 2, Application US/09364206
; Patent No. 6475752
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Tang, Y. Tom
; APPLICANT: Baugh, Mariah R.
; TITLE OF INVENTION: Mammalian Imidazoline Receptor
; FILE REFERENCE: PC-0006 US
; CURRENT APPLICATION NUMBER: US/09/364,206
; CURRENT FILING DATE: 1999-07-30
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: PERL Program
; SEQ ID NO 2
; LENGTH: 1504
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY:
; OTHER INFORMATION: 129581CD1
; PUBLICATION INFORMATION:
US-09-364-206-2

```

```

Query Match      6.7%; Score 98; DB 4; Length 1504;
Best Local Similarity 22.6%; Pred. No. 0.66;
Matches 65; Conservative 33; Mismatches 103; Indels 86; Gaps 14;

QY 30 EKEGQTVSGMFRAL-----GQGVSPATVPCISPELLAHLGQAMAHAP 73
   ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
Db 444 EKELDTVE-VLKATQKAKEVSKLSNPKKGGESRLSAAFCIRPSSPPTVAPASALP 502
   ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
QY 74 QPLLPMYRKLRFVSGSNAVPAPEESFEVWLEQATEIVKE-WPVTEAEKKRWLAESLRG- 131
   ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
Db 503 QPIL-----SNQIMFVQEEALASSLSDSLTPEHQPIAQG-----CSDSLES1 547
   ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
QY 132 PA-----LDLHVIQADNPSISVECLFAFKQVF---GSLSRRRTAQVRYLKYQEE 180
   ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
Db 548 PAGQAASDDLDRDVPAGVGAGSP-----EHAPEVQVVPVPGSGQIIFLPFTCIGYTATNQD- 601
   ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
QY 181 GEKVSAYVLRLETLRLKAVEKRAIPRIADQVRLEQVMAGATLNQMLWCLRELKDKQGP 240
   ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
Db 602 -----FIQLSTLIRQAIE-RQLP-----AWIEAANQREBQGG 633
   ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :

```

```

QY 241 PSFLELMKVIREEEERASPFENESIE-----EPEERDGYGRWNHGGDD 283
   ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
Db 634 EQGEE---EDEEEEDVAENRYFEMGPPDVEEEGGGQGEEREE 677
   ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :

RESULT 7
US-08-650-766-6
; Sequence 6, Application US/08650766D
; Patent No. 6015690
; GENERAL INFORMATION:
; APPLICANT: PILETZ, John E.
; APPLICANT: IVANOV, Tina R.
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A HUMAN IMIDAZOLINE RECEPTOR AND
; TITLE OF INVENTION: METHOD FOR CLONING THE SAME
; FILE REFERENCE: Corrected Sequence Listing
; Patent No. 6015690
; CURRENT APPLICATION NUMBER: US/08/650,766D
; CURRENT FILING DATE: 1996-05-20
; EARLIER APPLICATION NUMBER: US 60/012,600
; EARLIER FILING DATE: 1996-03-01
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 651
; TYPE: PRT
; ORGANISM: Homo sapiens
US-08-650-766-6

```

```

Query Match      6.7%; Score 97.5; DB 3; Length 651;
Best Local Similarity 23.0%; Pred. No. 0.21;
Matches 59; Conservative 30; Mismatches 99; Indels 69; Gaps 12;

QY 44 GQGVSPATVPCISPELLAHLGQAMAHAPQPLLPMYRKLRFVSGSNAVPAPEESFEVW 103
   ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
Db 61 GGEDSRLSAAFCIRPSSPPTVAPASASLPQIL-----SNQIMFVQEEALASS 110
   ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
QY 104 LEQATEIVKE-WPVTEAEKKRWLAESLRG-PA-----LDLHVIQADNPSISVECL 153
   ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
Db 111 LSSTDSLTPHQPIAQG-----CDSLESIPAGQAASDDLDRDVPAGVGAGSP-----EHA 160
   ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
QY 154 EAFKQVF---GSLSRRRTAQVRYLKYQEEGEKVSAYVLRLETLRLKAVEKRAIPRIAD 210
   ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
Db 161 EPEVQVVPVPGSGQIIFLPFTCIGYTATNQD-----FIQLSTLIRQAIE-RQLP----- 207
   ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
QY 211 QVRLEQVMAGATLNQMLWCLRELKDKQGPSPFLELMKVIREEEERASPFENESIE---- 266
   ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
Db 208 -----AWIEAANQREBQGGQGE-----EDEEEEDVAENRYFEMGPP 248
   ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
QY 267 EPEERDGYGRWNHGGDD 283
   ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
Db 249 DVEEEGGGQGEEREE 265
   ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :

```

```

RESULT 8
US-08-922-635-5
; Sequence 5, Application US/08922635A
; Patent No. 6033871
; GENERAL INFORMATION:
; APPLICANT: PILETZ, John E.
; APPLICANT: IVANOV, Tina R.
; TITLE OF INVENTION: DNA MOLECULES ENCODING IMIDALINE RECEPTIVE POLYPEPTIDES
; TITLE OF INVENTION: AND POLYPEPTIDES ENCODED THEREBY
; FILE REFERENCE: Corrected Sequence Listing
; Patent No. 6033871
; CURRENT APPLICATION NUMBER: US/08/922,635A
; CURRENT FILING DATE: 1997-09-03
; EARLIER APPLICATION NUMBER: 08/650,766
; EARLIER FILING DATE: 1996-05-20
; EARLIER APPLICATION NUMBER: 60/012,600
; EARLIER FILING DATE: 1996-03-01
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5

```

```

; LENGTH: 651
; TYPE: PRT
; ORGANISM: Homo sapiens
US-08-922-635-5

Query Match          6.7%; Score 97.5; DB 3; Length 651;
Best Local Similarity 23.0%; Pred. No. 0.21;
Matches 59; Conservative 30; Mismatches 99; Indels 69; Gaps 12;

QY  44  GQGVSPATVPCISPELLAHILGQAMAHAPOLIPMYRKLRVFSGSAVPAPERSEFVW 103
      : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db  61  GGEDSRLSAAPCIRPSSSPPTVAPASLPOFIL-----SNOGIMFQEEALASS 110
      : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
QY  104  LEQATEIVKE-WPYTEAEKKWLAESLRG-PA-----LQLMHIVQADNPISVSECL 153
      : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db  111  LSSTDSTUPBHQTAQG-----CSDLSIESIAGQAASDDLDDVPGAVGASD-----EHA 160
      : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
QY  154  EAFKQVF---GSLESRRTAQVRYLKTYOEQEKVSAYVLRLETLRLKAVERRAI 210
      : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db  161  EPEVQVVGSGQIIFLPTCIGYATNQD-----FIQLSTLIRQAIE-RQLP----- 207
      : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
QY  211  QVRLEQVMAGATLNMQLWCRLRELRKQGPSPFELMKVIRREBEERESAFNESTE---- 266
      : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db  208  -----AMIEAANQREEGQGGQEE-----EDEEESEEDVAENRYEMEGMP 248
      : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
QY  267  EPERDGYGRWNHEGDD 283
      : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db  249  DVEEEGGGGGEEEEEE 265
      : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :

```

```

RESULT 9
US-09-389-487-6
; Sequence 6, Application US/09389487
; Patent No. 6576742
; GENERAL INFORMATION:
; APPLICANT: PILETZ, John E.
; APPLICANT: IVANOV, Tina R.
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A HUMAN IMIDAZOLINE RECEPTOR AND
; TITLE OF INVENTION: METHOD FOR CLONING THE SAME
; FILE REFERENCE: Corrected Sequence Listing
; Patent No. 6576742
; CURRENT APPLICATION NUMBER: US/09/389,487
; CURRENT FILING DATE: 1999-09-03
; EARLIER APPLICATION NUMBER: US 08/650,766
; EARLIER FILING DATE: 1996-05-20
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 651
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-389-487-6

```

[illegible][illegible]

```

RESULT 11
US-09-264-512B-2
; Sequence 2, Application US/09264512B
; Patent No. 6610508
; GENERAL INFORMATION:
; APPLICANT: Hentze, Matthias W.
; TITLE OF INVENTION: De Gregorio, Ennio
; TITLE OF INVENTION: TRANSLATION DRIVER SYSTEM AND METHODS FOR USE THEREOF
; FILE REFERENCE: 9882-004
; CURRENT APPLICATION NUMBER: US/09/264,512B
; CURRENT FILING DATE: 1999-03-08
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 2
; LENGTH: 1560
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-264-512B-2
Query Match          6.4%; Score 94; DB 4; Length 1560;

```

Best Local Similarity 22.0%; Pred. No. 1.8;
Matches 82; Conservative 40; Mismatches 121; Indels 130; Gaps 16;

QY 14 TPNQTEFLERLNLFLEKEG-----QTSGMFRALQOEGVSPATVPCISPELLAHL 65
DB 410 SPAQEEEMEEEEE-----EEGEAGEAGEASEKG-----GEELLPPESTFI--PANLSQNL 459
QY 66 GQAMAHAPQLLPMRYKLRVFS----- 88
DB 460 EAAATQAVSVPRKRRIKELNKEAVGDLDDAFKEANPAVPEVNOPPAGSNPGPSE 519
QY 89 GSAVPAPEEESFEVWLEQATEI-----VKEW-PVTEAEKKRWLAESLRG 131
DB 520 GSGVPPRPEADETWDSKEDIHNAENIQGEQKYEKSDQWKPNNLEKKRYDRFLG 579
QY 132 -----PALDLWHIVQA-----DNPSISVECELEAFKQVFGSL----- 163
DB 580 QFTFASMQKEGEPGLPHISDVVLDKANKTLPRLDPTLRQINGCGPDTFSPANLGRITLS 639
QY 164 -----ESRTAQV-----RYLKVQEEGKVSAYVLRLETLRLKAVEKRAIP-- 205
DB 640 TRGPPRGPGGELPRGPQAGLGRPSQGPKEPKIATVMTEDIKLNKAEKAWKPSS 699
QY 206 RRIADQVRLQVWAGATNLNMLCRLRELKXQGPSPFLMLKVIRE-----EEEE----- 256
DB 700 KRTAADKDRGEEDADGSK-TQDLFRVRSLNKLTPQMFQQLMKQVTLADTTEERLKV 758
QY 257 -EASFENESTEEP 268
DB 759 IDLIFE-KAISEP 770

RESULT 12
US-08-556-419-25
; Sequence 25, Application US/08556419C
; Patent No. 6093549
; GENERAL INFORMATION:
; APPLICANT: Ross, Christopher
; APPLICANT: Li, Xiao-Jiang
; APPLICANT: Li, Shi-Hua
; APPLICANT: Sharp, Alan
; APPLICANT: Lananhan, Anthony
; APPLICANT: Worley, Paul
; APPLICANT: Snyder, Solomon
; TITLE OF INVENTION: Huntingtin-associated protein
; FILE REFERENCE: 01107.52271
; CURRENT APPLICATION NUMBER: US/08/556,419C
; CURRENT FILING DATE: 1995-11-09
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Fast-Seq for Windows Version 3.0
; SEQ ID NO 25
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Homo sapiens
US-08-556-419-25

Query Match 6.4%; Score 93.5; DB 3; Length 331;
Best Local Similarity 22.6%; Pred. No. 0.2;
Matches 53; Conservative 40; Mismatches 85; Indels 57; Gaps 11;

QY 54 PCTSPELLAH--LLQAMAHAPQLLPMRYKLRVFGSSAVPAPEESFEVWLEQATEIV 111
DB 44 PCDAKPLISQEAHLHQ--HHCPO--LEALQEKRLLL-----EENHQL-REASQ-- 88
QY 112 KENPVTEAEKKRWLAESLRGALDMLHIVQADNPSISVECELEAFKQVFGSLRRTAQV 171
DB 89 -----LDLLE-----DEQMLILECVQFSEASQMAELSEVLV 122
QY 172 RYLKTY---OECEKVSAYVLRLETLRL-----KAVEKRAIPRRIADQVRLQVWAG-- 220
DB 123 LRLENVERQOQEVARIQAQVKLQQRQCRMYGAETKQLQASEKEIQMLQEEETLPGF 182
QY 221 -ATLNQMLWCRRLRELKXQGPSPFLMLKVIREEEEEASFENESTEEPBERDGY 274

DB 183 QETLAELRTSLRRMISD--PVYFERNYEMPRGDTSSLRDYDFRYSDEQVRGF 235
US-07-853-913-2
; Sequence 2, Application US/07853913
; Patent No. 5338839
; GENERAL INFORMATION:
; APPLICANT: McKay, Ronald D.G.
; APPLICANT: Lendahl, Urban
; TITLE OF INVENTION: Nestin Expression As An Indicator of
; TITLE OF INVENTION: Neuroepithelial Tumors
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/853,913
; FILING DATE: 19920319
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/660,412
; FILING DATE: 22-FEB-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/603,803
; FILING DATE: 25-OCT-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/201,762
; FILING DATE: 02-JUN-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/180,548
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: MIT-4641AAAA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1805 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-853-913-2

Query Match 6.4%; Score 93; DB 1; Length 1805;
Best Local Similarity 22.8%; Pred. No. 2.9;
Matches 68; Conservative 35; Mismatches 137; Indels 58; Gaps 10;

QY 15 PNQDTEFLERLNLFLEKEGQTSGMFRALQOEGVSPATVPCISPELLAHLQAMAHAPQ 74
DB 664 PGADQMLEKL-----VKEDQSPRSPSEEDQACRFLQKENQBPGLGYEEAGQILE---- 716
QY 75 PLLPMRYKLRVFGSSAVPAPEESFEVWLEQATEIVKWPVTEAEK-----KRWLA 127
DB 717 -----RLIEKESQSLSRSPSEEDQACRFLQKENQBPGLGYEEADQMLRLIEKESQ 769
QY 128 SLRGPALDMLHIVQADNPSISVECELEAFKQVFGSLRRTAQVRLKTYQEEG----- 181
DB 770 SLKSPENORIGKPLERENQKSLRYLEENQETVPPLSRNQRPLRSLVEEERQIVKPL 829
QY 182 EKVSAYVLRLETLRLKAVEKRAIPRIADQVRLQVWAGATLNQMLWCRRLKDOG--- 238

Db 830 EKVS-----QDSGLSLAEENVQPIRYLEE---DCINKSLLEDKTHKSLGSLDRNGDS 880
QY 239 -----PPPSFLEMLKVIREEBEEAEFENESIEEPE--BRDGYGRWNHE 280
Db 881 IIIPOESETQVSLRPPEE--EDQIRVNHLEKESQFSRSSEEEQWMSLEGE-NHE 935

RESULT 14

US-09-252-991A-19048
; Sequence 19048, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19048
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19048

Query Match 6.3%; Score 92; DB 4; Length 736;
Best Local Similarity 22.9%; Pred. No. 0.95;
Matches 56; Conservative 33; Mismatches 98; Indels 58; Gaps 11;
QY 6 GYWKVIFKTPNDTEFLERLNLFLKEGQTVSGMFRALQGEVSP-----ATVPC 55
Db 68 GHWRLVEAPPSEEEK-----LEGYGLVPPFLPLPGDGSQPPAPSAASEAETPPDA 119
QY 56 ISPELLAHLLGQAMAHAPOLIPMYRKLVRP--SGSAVPAPEESFFVWLEQATEIVKE 113
Db 120 PAPSPALAEARQMGAEA---LPEKTAGAEFFEGEGSCRCRNDQDSALAPLRQ---VRD 171
QY 114 WPVTEAEKKR-----WLAESLRGPALDLMHIVQADNPSISVEECLEAFKQVF 160
Db 172 AGLGAEATKALANSRLDILGACQWEQEEELGG-----VLAQGVESAAGKAFATYLEAA 223
QY 161 GSLEGRRTAQVRY-LKTYQEEGE---KVSAYVLRLETL-----RKAVEKRAIPR-RIADQ 211
Db 224 ANFYSGRDEAEQGFALQDVSPWLKETAIVLQARTLLNAAQNAQNAFDDMGFPQLQNVDK 283
QY 212 VRIEQ 216
Db 284 ARLEQ 288

RESULT 15

US-08-056-200-94
; Sequence 94, Application US/08056200
; Patent No. 5616500
; GENERAL INFORMATION:
; APPLICANT: Steinert, Peter M.
; APPLICANT: Lee, Seung-Chul
; APPLICANT: Kim, In-Gyu
; APPLICANT: Chung, Soo-Il
; APPLICANT: Park, Sang-Chul
; TITLE OF INVENTION: Trichohyalin and Transglutaminase-3 and
; TITLE OF INVENTION: Methods of Using Same
; NUMBER OF SEQUENCES: 117
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive, Sixteenth Floor
; CITY: Newport Beach
; STATE: CA

; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/056,200
; FILING DATE: 30-APR-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Fedrick, Michael F.
; REGISTRATION NUMBER: 36,799
; REFERENCE/DOCKET NUMBER: NIH054.001A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (714) 760-0404
; TELEFAX: (714) 760-9502
; INFORMATION FOR SEQ ID NO: 94:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1898 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
US-08-056-200-94

Query Match 6.3%; Score 92; DB 1; Length 1898;
Best Local Similarity 26.1%; Pred. No. 4;
Matches 48; Conservative 24; Mismatches 82; Indels 30; Gaps 5;
QY 96 EESFEVWLEQATEIVKEWPVTEAKKRWLAESLRGPALDLMHIVQADNPSISVEECL-- 153
Db 500 EEERERWKLKEEERREQERREQLRREQERREQLRKEEERLQOQLRSEQQLRR 559
QY 154 ----EAPKQVFGSLESRRTAQVR---YLKTYQEEGEKVSAYVLRLETLRKAVEKRAIPR 207
Db 560 EOEERLEQLLKREERKLEQERREQLRKEQEE-----RDQLLKREERKQORLK 610
QY 208 IADQVRLQVLMAGATINQMLWCRLRELKDQGPSPFLELMKVIRREEEESFENESIEE 267
Db 611 REQERLEQLRKREEVERL---EQERRDE-----RLKREPEPEERREHLLKSEE 657
QY 268 PEER 271
Db 658 QEER 661

Search completed: September 21, 2004, 13:36:18
Job time : 22.0409 secs

Result No.	Query %			DB	ID	Description
	Score	Match	Length			
1	1459	99.8	283	3	AB12528	Human Ma4
2	1459	99.8	364	7	ADC08977	Onconeuro
3	755	51.6	149	3	AB12527	Human Ma3
4	628	43.0	353	4	AAW74701	Human mem
5	618.5	42.3	455	5	ABW05727	Human sig
6	618.5	42.3	463	3	AB12529	Human Ma5
7	618.5	42.3	463	3	AB12529	Human Ma5
8	597	40.8	351	4	AAW42315	Human ORF
9	596.5	40.8	351	4	AAE01340	Human pro
10	596.5	40.8	351	4	AAW74695	Human gen
11	596.5	40.8	351	4	AAW08664	Human mem
12	593	40.6	195	3	AAU08664	Human NOV
13	552	37.8	329	3	AB12526	Human Ma2
14	550	37.6	312	3	AB12525	Human Ma1
15	503.5	34.4	280	4	AB43023	Human ORF
16	480.5	32.9	237	4	AAE01336	Human gen
17	475.5	32.5	399	6	AAW94854	Human pro
18	475.5	32.5	399	6	ADA54410	Human pro
19	394	26.9	403	5	ABG99947	Human NOV
20	394	26.9	403	5	ABG97495	Human NOV
21	394	26.9	403	5	AAW51624	Human NOV
22	393	26.9	403	6	ABO14772	KIAA0883-
23	380.5	26.0	402	4	ABO14772	Novel hum
24	378.5	25.9	337	4	AAW60478	Novel hum
25	206.5	14.1	110	4	AAW01787	Human cel
						Human pro
						Human pol

[illegible]

RESULT 2	
ADC08977	
ID	ADC08977 standard; protein; 364 AA.
XX	
XX	AC
XX	ADC08977;
XX	
DT	18-DEC-2003 (first entry)
DE	Onconeural antigen Ma2 protein.
XX	
XX	Human; Ma2; onconeural; antigen; Alzheimer's disease;
KW	neurodegenerative disease; diagnosis; neuroprotective; gene therapy.
KW	

XX OS Homo sapiens.
XX WO2003073104-A2.
XX
XX 04-SEP-2003.
XX
XX 26-FEB-2003; 2003WO-EP001946.
XX
XX 26-FEB-2002; 2002EP-00004177.
XX 26-FEB-2002; 2002US-0359307P.
XX
XX (EVOT-) EVOTEC NEUROSCIENCES GMBH.
XX
XX Hipfel R, Von Der Kammer H, Pohlner J;
XX
XX WPI: 2003-721818/68.
XX GENBANK: O94959, KIAA0883.
XX
XX Diagnosing or prognosticating, or determining increased risk of
PT developing a neurodegenerative disease by determining level or activity
PT of a transcription or translation product of a gene coding for Ma
PT onconeural antigen.
XX
XX Disclosure; Fig 9; 51pp; English.
XX
XX The present sequence is the protein sequence of human onconeural
CC antigen Ma2. The invention discloses the detection and differential
CC expression and regulation of the Ma2 gene in specific brain regions of AD
CC patients. The Ma2 gene and its transcription and/or translation products
CC may have a causative role in the regional selective neuronal degeneration
CC typically observed in AD, or may confer a neuroprotective function to the
CC remaining nerve cells. Methods are claimed for diagnosing or

CC prognosticating a neurodegenerative disease, for monitoring the
CC progression of a neurodegenerative disease, and for evaluating treatment
CC of a neurodegenerative disease, especially AD, in a subject by
CC determining the level and/or activity of a transcription or translation
CC product of an Ma onconeural antigen gene, especially Ma2. Also claimed
CC are: a method for treating or preventing AD and related neurodegenerative
CC disorders using the Ma2 gene or its transcription or translation product;
CC a method of screening for modulating agents of neurodegenerative diseases;
CC ; and a recombinant non-human animal comprising an Ma2 gene sequence,
CC which is useful for screening, testing and validating candidate
CC diagnostic and therapeutic agents.

SQ		Sequence 364 AA;	
	Query Match	99.8%; Score 1459; DB 7; Length 364;	
	Best Local Similarity	99.6%; Pred. No. 5.5e-132;	
	Matches 282; Conservative 1;	Mismatches 0; Indels 0; Gaps 0;	
Qy	1	VQGKGVVWVFPTPNQDTFLRLNLELEKEGQTVSGMFRALGOEGVSPATVPICISPEL	60
Db	82	VQGKGVVWVFPTPNQDTFLRLNLELEKEGQTVSGMFRALGOEGVSPATVPICISPEL	141
Qy	61	LAHLIGQAMAHAPQLPMRYRKURVFGSVAVPAPEERSFEVWLSCATEIVKWPVTAE	120
Db	142	LAHLIGQAMAHAPQLPMRYRKURVFGSVAVPAPEERSFEVWLSCATEIVKWPVTAE	201
Qy	121	KRWLAESLRGPALDLMIHVOADNPISIVEECLEAFKVFGSLERRTAQVRYLKTYQE	180
Db	202	KRWLAESLRGPALDLMIHVOADNPISIVEECLEAFKVFGSLERRTAQVRYLKTYQE	261
Qy	181	GKVSAYVRLRLETLRLRAVEKRAIPRRITADQVRLEQVMAGATLNQMLWCRLRELKDQGGP	240
Db	262	GKVSAYVRLRLETLRLRAVEKRAIPRRITADQVRLEQVMAGATLNQMLWCRLRELKDQGGP	321
Qy	241	PSFTLMKVIRREBEESA FNESIEEPEDRGYGRWNHEGD	283
Db	322	PSFTLMKVIRREBEESA FNESIEEPEDRGYGRWNHEGD	364

RESULT 3	
AAB12527	
ID	AAB12527 standard; protein; 149 AA.
XX	
AC	AAB12527;
XX	
DT	02-NOV-2000 (first entry)
XX	
DE	Human Ma3 protein SEQ ID NO:9.
XX	
KW	Ma1; Ma2; Ma3; Ma4; Ma5; paraneoplastic limbic encephalitis; diagnosis;
KW	paraneoplastic syndrome; neoplasm; colon cancer; brainstem encephalitis;
KW	breast cancer; parotid gland cancer; lung cancer; testicular cancer;
KW	germ-cell tumour.
XX	
OS	Homo sapiens.
XX	
PN	JP2000146982-A.
XX	
PD	26-MAY-2000.
XX	
PF	10-NOV-1999; 99JP-00320171.
XX	
PR	10-NOV-1998; 98US-00189527.
XX	
PA	(SLOK) SLOAN KETTERING INST CANCER RES.
XX	
DR	WPI; 2000-468119/41.
DR	N-PSDB; AAA60835.
XX	
PT	Diagnosing paraneoplastic syndrome e.g. paraneoplastic limbic
PT	encephalitis or neoplasm e.g. colon cancer comprising assessing a test
PT	sample for the presence or absence of antibodies to a Ma family
PT	polypeptide.

XX PS Claim 48; Fig 6; 27pp; Japanese.

XX CC The present invention describes a method for diagnosing a paraneoplastic

CC syndrome or neoplasm. The method comprises assessing a test sample for

CC the presence or absence of antibodies to a Ma family polypeptide (I). The

CC method is used to diagnose a paraneoplastic syndrome especially

CC paraneoplastic limbic encephalitis and/or brainstem encephalitis or

CC neoplasm in an individual. The method diagnoses the neoplasm by assessing

CC antibodies to (I) preferably Ma1, which is indicative presence of breast

CC cancer, colon cancer, parotid gland cancer, lung cancer, testicular

CC cancer and germ-cell tumours or Ma2, which is indicative of testicular

CC cancer, germ-cell tumour, and lung cancer. The present sequence is the

XX Ma3 protein as given in the present invention

XX Sequence 149 AA;

Query Match 51.6%; Score 755; DB 3; Length 149;

Best Local Similarity 98.7%; Pred. No. 1.4e-64;

Matches 147; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 135 DLHIVQADNPISVEECLEAFKQVFGSLESPRTAQVRLKTYQEGKVSAYVRLLETL 194

Db 1 DLHIVQADNPISVEECLEAFKQVFGSLESPRTAQVRLKTYQEGKVSAYVRLLETL 60

QY 195 LRKAVEKRAIPRRIADQVRLQVWAGATLNQMLWCLRELKXQDQPPPSFLELMKVIREEE 254

Db 61 LRRAVEKRAIPRRIADQVRLQVWAGATLNQMLWCLRELKXQDQPPPSFLELMKVIREEE 120

QY 255 EEEASFENESIEEPEDRGYGRWNHGGDD 283

Db 121 EEEASFENESIEEPEDRGYGRWNHGGDD 149

RESULT 4

AA74701

AA74701 standard; protein; 353 AA.

AA74701;

12-JUN-2001 (first entry)

Human membrane associated protein MEMAP-7.

Human; membrane associated protein; MEMAP; diagnosis; cytostatic;

KW antiinflammatory; anticonvulsant; immunosuppressive; antidiarrheic;

KW antiarteriosclerotic; gene therapy; cell proliferative disorder;

KW autoimmune disorder; inflammatory disorder; neurological disorder;

KW gastrointestinal disorder; cancer; inflammation; atherosclerosis;

KW epilepsy; diarrhoea.

OS Homo sapiens.

XX W0200112662-A2.

XX 22-FEB-2001.

XX 14-AUG-2000; 2000WO-US022315.

XX 17-AUG-1999; 99US-0149641P.

XX 09-NOV-1999; 99US-0164203P.

XX (INCY-) INCYTE GENOMICS INC.

XX Lal P, Yue H, Tang YT, Bandman O, Burford N, Azimzai Y;

PI Baughn MR, Lu DAM, Patterson C;

XX WPI; 2001-169860/17.

XX N-PSDB; AAF81747.

XX Isolated polypeptide with a human membrane associated protein sequence is

PT useful for the diagnosis, prevention and treatment of cell proliferative,

PT autoimmune/inflammatory, neurological and gastrointestinal disorders.

XX PS Claim 1; Page 119-120; 173pp; English.

XX CC AAF81741 to AAF81777 encode the human membrane associated proteins

CC (MEMAP) given in AAB74695 to AAB74731. MEMAPs have cytostatic,

CC antiinflammatory, anticonvulsant, immunosuppressive, antidiarrheic and

CC antiarteriosclerotic activities, which can be used in gene therapy.

CC MEMAPs and agonist of MEMAPs can be used to treat a disease or condition

CC associated with decreased expression of functional MEMAP and antagonists

CC of MEMAP are used to treat a disease or condition associated with

CC overexpression of functional MEMAP. These disorders include cell

CC proliferative, autoimmune/inflammatory, neurological and gastrointestinal

CC disorders. The MEMAP polynucleotides and proteins are also used for the

CC diagnosis of these disorders. Specific examples of these disorders

CC include cancer, inflammation, atherosclerosis, epilepsy and diarrhoea.

CC MEMAP proteins can be used to screen for compounds which specifically

CC bind MEMAP including antibodies, oligonucleotides, proteins and small

CC molecules. MEMAP polynucleotides can be used to prepare transgenic

CC animals which can be studied to provide information concerning human

CC disease. Anti-MEMAP antibodies are useful in immunoassays for the

CC detection of MEMAP protein and can be used as antagonists to treat or

CC prevent a disorder associated with MEMAP. Polynucleotides encoding MEMAP

CC can be delivered to target cells with genetic abnormalities with respect

CC to the expression of MEMAP to treat or prevent a disorder associated with

CC MEMAP

XX Sequence 353 AA;

Query Match 43.0%; Score 628; DB 4; Length 353;

Best Local Similarity 47.9%; Pred. No. 9e-52;

Matches 128; Conservative 52; Mismatches 75; Indels 12; Gaps 4;

QY 3 KGKGVKVIKTPNODTEFLERLNLFLKEGQTVSGMFRALQGVSPATVPCISPPELLA 62

Db 83 KGKGVKVLKFPPTSDAEFLERLHLFLAREGVTQDVARVLGFQNPPTP---GPMPA 138

QY 63 HLLGQAMAHAPQPLL-PMRYKRLVFGSGAVPAPDEESFEVWLEQATEIVKEWPTEAEK 121

Db 139 EMLNYILDNVIQPLVESITWKRLTLFGGRDIPFGGEETFDPLWLEHTNVELEEQVSDVEK 198

QY 122 KRWLAEISLRGPDALDHIVQADNPISVEECLEAFKQVFGSLESPRTAQVRLKTYQEG 181

Db 199 RRLMESIRGPAADVIRILKSNPAITTAECLEQVFGSVESRRDQIRKLTNYQNGP 258

QY 182 EKVSAVIRLETLRKAVEKRAIPRRIADQVRLQVWAGATLNQMLWCLRELKXQD 237

Db 259 EKLSAVIRLEPLLKQVVEKGAIDKDNVQARLEQVIAGANHSQAIRQLWL---TGAGE 315

QY 238 GPPPSFLELMKVIREEEESAFENES 264

Db 316 GPAPNLFQQLVQIREEEAKKEEEAE 342

RESULT 5

ABB05727

ID ABB05727 standard; protein; 455 AA.

XX ABB05727;

XX 30-APR-2002 (first entry)

XX Human signal transduction protein clone tes3_5k22.

XX Human; foetal brain; foetal kidney; melanoma; testis; amygdala;

XX gene therapy.

XX Homo sapiens.

XX W02001198454-A2.

XX 27-DEC-2001.

XX 25-APR-2001; 2001WO-IB002050.

KW	antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
KW	neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW	cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW	cholesterol ester storage; systemic lupus erythematosus; infection;
KW	severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW	allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW	bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW	thrombosis; contraceptive.
XX	
XX	Homo sapiens.
OS	
PN	WO200058473-A2.
PN	
XX	
XX	05-OCT-2000.
PD	
XX	
XX	31-MAR-2000; 2000WO-US008621.
XX	
XX	
PR	31-MAR-1999; 99US-01276707P.
PR	02-APR-1999; 99US-0127636P.
PR	05-APR-1999; 99US-0127728P.
PR	30-MAR-2000; 2000US-00540763.
XX	
XX	(CURA-) CURAGEN CORP.
PA	
XX	
XX	Shimkets RA, Leach M;
PI	
XX	
DR	WPI; 2000-602362/57.
DR	N-PSDB; AAC76524.
XX	
XX	
PT	Novel nucleic acids and peptides derived from open reading frame X,
PT	useful for treating e.g cancers, proliferative disorders,
PT	neurodegenerative disorders and cardiovascular disease.
XX	
PS	Claim 11; Page 3345-3347; 5507pp; English.

AAC774446 to AAC77606 encode the proteins given in ABA403237 to ABA43397, which represent the human ORFX open reading frames 1 to 3161. The ORFX sequences have activities such as: cytostatic; hepatotropic; vulnery; antispasmodic; antiparkinsonian; nootropic; neuroprotective; osteopathic; anticoagulant; antiarthritic; immunosuppressant; immunostimulant; cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive; dermatological; immunosuppressive; antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic; antithyroid; and antianaemic. The sequences can be used for determining the presence of or predisposition to, or preventing or treating pathological conditions associated with an ORFX-associated disorder. The nucleic acids can be used to express ORFX proteins in gene therapy vectors. The proteins and nucleic acids may be used to treat cancers, proliferative disorders, neurodegenerative disorders, osteoarthritis, graft vs host disease, cardiovascular disease, diabetes mellitus, hypertension, hypothyroidism, cholesterol ester storage, systemic lupus erythematosus, severe combined immunodeficiency (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to enhance coagulation; to inhibit thrombosis; and as a contraceptive

	Query Match	42.3%;	Score 618.5;	DB 3;	Length 463;
	Best Local Similarity	50.2%;	Pred. No. 1.1e-50;		
	Matches 135; Conservative	44;	Mismatches 85;	Indels 5;	Gaps 3;
QY	1	VQGKGGVWKVFIFKTPNODTEFLERLNLFLEKEGQTSGMFRALCOEGVGSPATVFCISPEL	60		
	:	:	:	:	:
Db	82	IPKGCGFWFIVKPRNSDGFNLRLNFLSEERTVSDMRVLGSDTNCAPRYVTISPEF	141		
	:	:	:	:	:
QY	61	LAHLGQAAMAHAPQLL--PMRYRKLRYFGSSAVDPAPEESFEVWLEQATEIVKSWPVEA	119		
	:	:	:	:	:
Db	142	WT--WAQTLGAAYQPILLEQMLYRELRYFVSGNTISIPCALAFDAWLEHTMLQMWQVPEG	199		
	:	:	:	:	:
QY	120	EKKRWLAESIRGPALDLMHIVQADNPISISVEBCLAEAKQVFGSLESRRTAQVRKYLTQYE	179		
	:	:	:	:	:
Db	200	EKKRRUMELCRGPALQVSGLRASNASITVEECLAALQQVFGPVESHKAQVKLCCKAYQE	259		
	:	:	:	:	:

[illegible]

RESULT 8	
AAO16179	
ID	AAO16179 standard; protein; 452 AA.
XX	
XX	AAO16179;
XX	
DT	28-MAR-2003 (first entry)
XX	
XX	Human protein #5.
DE	
XX	
XX	Human; vaccine; adult whole brain; foetal whole brain; tonsil;
KW	adult hippocampus; disease-associated SNP analysis; knockout mouse;
KW	disease model mouse; cancer; neurological disorder.

OS	Homo sapiens.
XX	
XX	
PN	WO200299103-A1.
XX	
PD	
XX	12-DEC-2002.
XX	
XX	
PF	27-MAY-2002; 2002WO-JP005134.
XX	
PR	04-JUN-2001; 2001JP-00168370.
PR	16-AUG-2001; 2001JP-00246915.
PR	

PA	(KAZU-) KAZUSA DNA RES INST FOUND.
PA	(PROT-) PROTEIN EXPRESS CO LTD.
XX	
PI	Ohara O, Nagase T, Nakajima D;
XX	
DR	WPI: 2003-140622/13.
DR	N-PSDB; AAL51207.
XX	
PT	DNA preferentially expressed in human adult and fetal brain tissue
PT	for diagnosis, treatment and analysis of cancer and mental disorder
XX	
PS	Claim 1; Page 56-60; 73pp; Japanese.

The invention comprises the amino acid and coding sequences of seven human proteins that are preferentially expressed in adult whole brain, foetal whole brain, tonsil and adult hippocampus tissue. The DNA sequences are useful for the analysis of disease-associated single nucleotide polymorphisms and the production of knockout human disease model mice. The DNA and protein sequences of the invention are useful for the prevention (vaccine) and treatment of cancer and neurological disorders. The present amino acid sequence represents a human protein of the invention.

[illegible]

Db 194 QVSEKRRRLLESRLGSPALSMVRLQNNDSITVEQCLDALKQIFGKEDFRASQREFL 253
QY 175 KTYQEGEKVSAYVLRLETLTKARVEKRAIPRIADQVRLSQVAGATVNLQMLWCRRLREL 234
Db 254 QTSPIKIGKSVFTLLRLPLQLQKAVHKSPLSVSRSTMIRLKHLLARVAMTPALRGKLELL 313
QY 235 KDGPPPPFLELMKVIREEE---FEASFENESIEPEERQGYGR 276
Db 314 DQGGPPNFLELMKLIRDEEWENTAVMKNK-----EKPSGRGR 353

RESULT 9
AAE01340
ID AAE01340 standard; protein; 351 AA.
AC AAE01340;
XX 17-JUL-2001 (first entry)
DE Human gene 22 encoded secreted protein fragment, SEQ ID NO:205.
XX Human; secreted protein; proliferative disorder; cancer; tumour;
KW foetal abnormality; developmental abnormality; haematopoietic disorder;
KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
KW inflammation; allergy; neurological disorder; Alzheimer's disease;
KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;
KW skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
KW cardiovascular disorder; angiotensin-related disorder;
KW gastrointestinal disorder; pregnancy-related disorder;
KW endocrine disorder; infection; wound healing; vulnerability; cell culture;
KW chemotaxis; food additive; gene therapy; binding partner identification.
XX Homo sapiens.
XX WO200134769-A2.
XX 17-MAY-2001.
XX 01-NOV-2000; 2000WO-US030040.
XX 05-NOV-1999; 99US-0163580P.
XX 30-JUN-2000; 2000US-0215130P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Ruben SM, Komatsoulis GA, Wei P, Baker KP, Fiscella M;
XX WPI; 2001-308781/32.
XX New isolated nucleic acid molecule encoding a human secreted protein is
XX used in preventing, treating or ameliorating a medical condition.
XX PS Disclosure; Page 46; 519pp; English.
XX AAD05121-AAD05203 represent cDNAs corresponding to 24 human secreted
XX protein genes, and AAE01232-AAE01311 represent the proteins they encode.
XX AAE01312-AAE01340 represent human secreted protein variants or fragments.
XX The secreted proteins and their genes are useful for preventing, treating
XX or ameliorating medical conditions, e.g., by protein or gene therapy.
XX Pathological conditions can be diagnosed by determining the amount of the
XX new protein in a sample or by determining the presence of mutations in
XX the new genes. Specific uses are described for each of the 24 genes,
XX based on the tissues in which they are most highly expressed, and include
XX developing products for the diagnosis or treatment of proliferative
XX disorders, cancer, tumours, foetal and developmental abnormalities,
XX haematopoietic disorders, diseases of the immune system, AIDS, autoimmune
XX diseases (e.g., rheumatoid arthritis), inflammation, allergies,
XX neurological disorders (e.g., Alzheimer's disease, Parkinson's disease),
XX cognitive disorders, schizophrenia, asthma, skin disorders (e.g.,
XX psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders,
XX angiogenic disorders, kidney disorders, gastrointestinal disorders,
XX pregnancy-related disorders, endocrine disorders, and infections.

CC proteins can also be used to aid wound healing and epithelial cell
CC proliferation, to prevent skin aging due to sunburn, to maintain organs
CC before transplantation, for supporting cell culture of primary tissues,
CC to regenerate tissues, to identify their cognate ligands or binding
CC partners, and in chemotaxis, and can be used as a food additive or
CC preservative to modify storage properties. Antibodies specific for a
CC protein of the invention can be used in alleviating symptoms associated
CC with the disorders mentioned above, and in diagnostic immunoassays e.g.,
CC radioimmunoassay or enzyme linked immunosorbent assay (ELISA). The
CC present sequence represents a human secreted protein fragment referred to
CC in the disclosure of the invention
XX SQ Sequence 351 AA;
Query Match 40.8%; Score 596.5; DB 4; Length 351;
Best Local Similarity 48.3%; Pred. No. 9.8e-49;
Matches 128; Conservative 49; Mismatches 75; Indels 13; Gaps 5;
QY 1 VOGKGVWVKVIFKTPNQDTEFLERLNLFLKEGQTVSGMFRALGQGVSPATVPCISPEL 60
Db 82 IPGKGGIWRVIFKPPDPDNTFLSRLNEFLAGEGTMVGLSRALGHENGSLDPEQGMIPEM 141
QY 61 LAHLGQAMAHAPQLLP-MRYRKLRFVSGSAVPAPEESFEVLEQATEIVKEWPVTEA 119
Db 142 WAPMLAQAL-EALQALQCLTKYKLRVSGRSPGPEEGFGRWFHTOMIKAMQVDPV 200
QY 120 EKRWLABSLRGPALDLMHIVQADNPISVVECLFAFKQVFGSLSRRTAQRVYLYKTYOE 179
Db 201 EKRRLLLESRLGPDLDVIRLVKINPLITVDECLQALEVFGVTDNPRELQVKYLTYYQK 260
QY 180 EGEKVSAYVLRLETLTKARVEKRAIPRIADQVRLSQVAGATVNLQMLWCRRLRELKD 236
Db 261 DEEKL SAYVLRLEPLQLQKAVHQAERDAVNOARLDQVIAGAVHKTIRREL-----NLPE 315
QY 237 QGPPSPFLELMKVIRE---EEEEEA 258
Db 316 DGPAPGFLQLLVLIKDYEAEEEEEA 340

RESULT 10
AAB74695
ID AAB74695 standard; protein; 351 AA.
XX AAB74695;
XX 12-JUN-2001 (first entry)
XX Human membrane associated protein MEMAP-1.
XX Human; membrane associated protein; MEMAP; diagnosis; cytostatic;
KW antiinflammatory; anticonvulsant; immunosuppressive; antidiarrheic;
KW antiarteriosclerotic; gene therapy; cell proliferative disorder;
KW autoimmune disorder; inflammatory disorder; neurological disorder;
KW gastrointestinal disorder; cancer; inflammation; atherosclerosis;
KW epilepsy; diarrhoea.
XX Homo sapiens.
XX WO200112662-A2.
XX 22-FEB-2001.
XX 14-AUG-2000; 2000WO-US022315.
XX 17-AUG-1999; 99US-0149641P.
XX 09-NOV-1999; 99US-0164203P.
XX (INCY-) INCYTE GENOMICS INC.
XX Lal P, Yue H, Tang YT, Bandman O, Burford N, Azimzai Y;
XX Baughn MR, Lu DAM, Patterson C;
XX WPI; 2001-168860/17.


```

QY 120 EKKWLABSLRGPDALDLMHIIVQADNPSISVECELEAFKQVFGSLESRRTAQVRYLKYOE 179
DB 201 EKRRLESLSRGPDALDVRVLKINPLITVDECLQALEVFGVTDPNRELQVKYLYTTQK 260

QY 180 EGEKVSAYVRLLETLLRKAVERKRAIPRIADQVRLQVWAGA---TLNQLMWCRLRELKD 236
DB 261 DEEKUSAVVRLLEPLQLKLVQGAERDAVNOARLDQVIAGAVHKTIRREL-----NLPE 315

QY 237 QGPPSFLELMKVIRE---EEEEEA 258
DB 316 DGPAPGFLQLLVLIKDYEAABEEEA 340

RESULT 12
AAB12526
ID AAB12526 standard; protein; 195 AA.
XX
AC AAB12526;
XX
DT 02-NOV-2000 (first entry)
XX
DE Human Ma2 protein SEQ ID NO:7.
XX
KW Ma1; Ma2; Ma3; Ma4; Ma5; paraneoplastic limbic encephalitis; diagnosis;
paraneoplastic syndrome; neoplasm; colon cancer; brainstem encephalitis;
breast cancer; parotid gland cancer; lung cancer; testicular cancer;
germ-cell tumour.
XX
OS Homo sapiens.
XX
PN JP2000146982-A.
XX
PD 26-MAY-2000.
XX
PF 10-NOV-1999; 99JP-00320171.
XX
PR 10-NOV-1998; 98US-00189527.
XX
PA (SLOK ) SLOAN KETTERING INST CANCER RES.
XX
WPI; 2000-468119/41.
DR N-PSDB; AAB60834.
XX
Diagnosing paraneoplastic syndrome e.g. paraneoplastic limbic
encephalitis or neoplasm e.g. colon cancer comprising assessing a test
sample for the presence or absence of antibodies to a Ma family
polypeptide.
XX
Claim 48; Fig 2; 27pp; Japanese.
XX
The present invention describes a method for diagnosing a paraneoplastic
syndrome or neoplasm. The method comprises assessing a test sample for
the presence or absence of antibodies to a Ma family polypeptide (I). The
method is used to diagnose a paraneoplastic syndrome especially
paraneoplastic limbic encephalitis and/or brainstem encephalitis or
neoplasm in an individual. The method diagnoses the neoplasm by assessing
antibodies to (i) preferably Ma1, which is indicative presence of breast
cancer, colon cancer, parotid gland cancer, lung cancer, testicular
cancer, and germ-cell tumours or Ma2, which is indicative of testicular
cancer, germ-cell tumour, and lung cancer. The present sequence is the
Ma2 protein as given in the present invention
XX
Sequence 195 AA;
Query Match 40.6%; Score 593; DB 3; Length 195;
Best Local Similarity 98.3%; Pred. No. 9.2e-49;
Matches 113; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VQGGGKWKVIFKTPNQDTEFLERLNLFLKEGQTVSGMFRALGOEGVSPATVPCISP 60
DB 81 VQGGGKWKVIFKTPNQDTEFLERLNLFLKEGQTVSGMFRALGOEGVSPATVPCISP 140

```

```

QY 61 LAHLIGQAMAHAPQLLPMRYRKLVRFGSVAVPAPERSEFEVWLQATEIVKEWP 115
DB 141 LAHLIGQAMAHAPQLLPMRYRKLVRFGSVAVPAPERSEFEVWLQATEIVKEWP 195

RESULT 13
AAB12525
ID AAB12525 standard; protein; 329 AA.
XX
AC AAB12525;
XX
DT 02-NOV-2000 (first entry)
XX
DE Human Ma1 protein SEQ ID NO:4.
XX
KW Ma1; Ma2; Ma3; Ma4; Ma5; paraneoplastic limbic encephalitis; diagnosis;
paraneoplastic syndrome; neoplasm; colon cancer; brainstem encephalitis;
breast cancer; parotid gland cancer; lung cancer; testicular cancer;
germ-cell tumour.
XX
OS Homo sapiens.
XX
PN JP2000146982-A.
XX
PD 26-MAY-2000.
XX
PF 10-NOV-1999; 99JP-00320171.
XX
PR 10-NOV-1998; 98US-00189527.
XX
PA (SLOK ) SLOAN KETTERING INST CANCER RES.
XX
WPI; 2000-468119/41.
DR N-PSDB; AAB60833.
XX
Diagnosing paraneoplastic syndrome e.g. paraneoplastic limbic
encephalitis or neoplasm e.g. colon cancer comprising assessing a test
sample for the presence or absence of antibodies to a Ma family
polypeptide.
XX
Claim 48; Fig 1; 27pp; Japanese.
XX
The present invention describes a method for diagnosing a paraneoplastic
syndrome or neoplasm. The method comprises assessing a test sample for
the presence or absence of antibodies to a Ma family polypeptide (I). The
method is used to diagnose a paraneoplastic syndrome especially
paraneoplastic limbic encephalitis and/or brainstem encephalitis or
neoplasm in an individual. The method diagnoses the neoplasm by assessing
antibodies to (i) preferably Ma1, which is indicative presence of breast
cancer, colon cancer, parotid gland cancer, lung cancer, testicular
cancer, and germ-cell tumours or Ma2, which is indicative of testicular
cancer, germ-cell tumour, and lung cancer. The present sequence is the
Ma1 protein as given in the present invention
XX
Sequence 329 AA;
Query Match 37.8%; Score 552; DB 3; Length 329;
Best Local Similarity 46.8%; Pred. No. 1.8e-44;
Matches 117; Conservative 45; Mismatches 74; Indels 14; Gaps 5;

QY 3 GKGGVWKVIFKTPNQDTEFLERLNLFLKEGQTVSGMFRALGOEGVSPATVPCISP 62
DB 83 GKGGVWKVIFKTPNQDTEFLERLNLFLKEGQTVSGMFRALGOEGVSPATVPCISP 138

QY 63 HLIGQAMAHAPQLL-PMRYRKLVRFGSVAVPAPERSEFEVWLQATEIVKEWPTEAEK 121
DB 139 EMLNYIDNVIQPLVESIWKRLTLFSGKGHPRAWRGNFDPLWLEHTNEVLEEQVSDVEK 198

QY 122 KRWLAESLRGPDALDLMHIIVQADNPSISVECELEAFKQVFGSLESRRTAQVRYLKYOE 181
DB 199 TRRLMESLRGPAADVIRILKSNPNPAITTAELCLALEHVFSGVSSRDAQIKFLNTYQNG 258

QY 182 EKVSAYVRLLETLLRKAVERKRAIPRIADQVRLQVWAGA---TLNQLMWCRLRELKD 236

```

Db 259 EKLSAYVIRLEPQKVEKGAIDKDNVQARLEQVIAGNHSIAIRQLWLTGARE--- 315
Qy 237 QGPPPSFLEL 246
Db 316 -GFGPKPLSV 324
RESULT 14
AAB43023
ID AAB43023 standard; protein; 312 AA.
XX AAB43023;
XX
XX
XX 08-FEB-2001 (first entry)
XX Human ORFX ORF2787 polypeptide sequence SEQ ID NO:5574.
XX Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW vulnery; antiposoriatic; antiparkinsonian; nootropic; neuroprotective;
KW immunostimulant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW hypotensive; dermatological; coagulant; vasotropic; antidiabetic;
KW antiviral; antibacterial; immunosuppressive; antiinflammatory;
KW antianemic; gene therapy; cancer; proliferative disorder; antithyroid;
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW thrombosis; contraceptive.
XX
XX Homo sapiens.
XX
XX WO200058473-A2.
XX
XX 05-OCT-2000.
XX
XX 31-MAR-2000; 2000WO-US008621.
XX
XX 31-MAR-1999; 99US-0127607P.
XX 02-APR-1999; 99US-0127636P.
XX 05-APR-1999; 99US-0127728P.
XX 30-MAR-2000; 2000US-00540763.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Shimkets RA, Leach M;
XX
XX WPI; 2000-602362/57.
XX N-PSDB; AAC77232.
XX
XX Novel nucleic acids and peptides derived from open reading frame X,
XX useful for treating e.g. cancers, proliferative disorders,
XX neurodegenerative disorders and cardiovascular disease.
XX
XX Claim 11; Page 4759; 5507pp; English.
XX
XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
XX which represent the human ORFX open reading frames 1 to 3161. The ORFX
XX sequences have activities such as: cytostatic; hepatotropic; vulnery;
XX antiposoriatic; antiparkinsonian; nootropic; neuroprotective; osteopathic;
XX anticonvulsant; antiarthritic; immunosuppressant; immunostimulant;
XX cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;
XX dermatological; immunosuppressive; antiinflammatory; antibacterial;
XX antiviral; antifungal; antirheumatic; antithyroid; and antianemic. The
XX sequences can be used for determining the presence of or predisposition
XX to, or preventing or treating pathological conditions associated with an
XX ORFX-associated disorder. The nucleic acids can be used to express ORFX
XX proteins in gene therapy vectors. The proteins and nucleic acids may be
XX used to treat cancers, proliferative disorders, neurodegenerative
XX disorders, osteoarthritis, graft vs host disease, cardiovascular disease,

CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester
CC storage, systemic lupus erythematosus, severe combined immunodeficiency
CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune
CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and
CC cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to
CC enhance coagulation; to inhibit thrombosis; and as a contraceptive
XX
XX Sequence 312 AA;
Query Match 37.6%; Score 550; DB 3; Length 312;
Best Local Similarity 50.9%; Pred. No. 2.5e-44;
Matches 113; Conservative 41; Mismatches 66; Indels 2; Gaps 2;
Qy 1 VQKGGVWKVIFKTPNQDTEFLERLNFLEKGGQTVSGMFRALGQEGVSPATVPCISPEL 60
Db 82 IPGGGIWRVIFKPPDPDNTFLSLRNLFLAGMTVGSLRGLHENGSLDPEQGMIPDM 141
Qy 61 LAHLGQAMAHAPQPLLP-MRYKLRVFGSAPPAPEEPEVWLEQATEIVKEWPVTEA 119
Db 142 WAPMLAQAL-EALQALQCLYKLRVSGRESPEEGEEFGRMWPHHTOMIKAWQVDPV 200
Qy 120 EKRWLAESLRGPAALDMHIVQADNPISVBECLEAFKQVFGSLRRRTAQRVYLKTYQE 179
Db 201 EKRRRLLESIRGPAALDVIRVLKINNPLITVDECLQALEEVGVDNPRELQVLYTTQK 260
Qy 180 EGEKUSAVVLETLRLKAVKRAIPRIADQVRLEQVMAGA 221
Db 261 DEEKL SAYVIRLEPQKVEKGAIDKDNVQARLEQVIAGNHSIAIRQLWLTGARE 302
RESULT 15
AAE01336
ID AAE01336 standard; protein; 280 AA.
XX
XX AAE01336;
XX
XX 17-JUL-2001 (first entry)
XX
XX Human gene 22 encoded secreted protein fragment, SEQ ID NO:201.
XX Human; secreted protein; proliferative disorder; cancer; tumour;
KW fetal abnormality; developmental abnormality; haematopoietic disorder;
KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
KW inflammation; allergy; neurological disorder; Alzheimer's disease;
KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;
KW skin disorder; psoriasis; sepsis; diabetes; kidney disorder;
KW cardiovascular disorder; angiogenic disorder; kidney disorder;
KW gastrointestinal disorder; pregnancy-related disorder;
KW endocrine disorder; infection; wound healing; vulnery; cell culture;
KW chemotaxis; food additive; gene therapy; binding partner identification.
XX
XX Homo sapiens.
XX
XX WO200134769-A2.
XX
XX 17-MAY-2001.
XX
XX 01-NOV-2000; 2000WO-US030040.
XX
XX 05-NOV-1999; 99US-0163580P.
XX 30-JUN-2000; 2000US-0215130P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Ruben SM, Komatsoulis GA, Wei P, Baker KP, Fiscella M;
XX WPI; 2001-308781/32.
XX
XX New isolated nucleic acid molecule encoding a human secreted protein is
XX used in preventing, treating or ameliorating a medical condition.
XX
XX Disclosure; Page 46; 519pp; English.
XX

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 13:20:22 ; Search time 14.7398 Seconds
(without alignments)
1272.567 Million cell updates/sec

Title: US-10-037-860-7
Perfect score: 996
Sequence: 1 PIALLEDWCRIMSDQKSL.....EESFEVLEQATEIVKEWP 195

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

PIR 78:*
1: Pirl:*
2: pirl2:*
3: pirl3:*
4: pirl4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	102	10.2	608	T03476	conserved hypotet
2	95	9.5	341	C72403	hypothetical prote
3	87.5	8.8	1684	T02632	hypothetical prote
4	85	8.5	751	F83080	hypothetical prote
5	84.5	8.5	3587	I40486	surfactin synthet
6	83.5	8.4	623	I64034	hypothetical prote
7	83	8.3	1132	T03668	phytochrome B - co
8	83	8.3	1295	T21720	hypothetical prote
9	83	8.3	1528	T06694	hypothetical prote
10	81.5	8.2	217	F69823	conserved hypotet
11	81.5	8.2	293	F69188	ethylene-inducible
12	81.5	8.2	435	S18609	glutamate dehydrog
13	81	8.1	330	G90256	DNA primase, proba
14	81	8.1	390	A82576	hypothetical prote
15	81	8.1	457	I5976	dihydroliipoamide S
16	81	8.1	502	T05246	cytochrome P450 mo
17	81	8.1	555	S21766	dihydroliipoamide S
18	81	8.1	609	A43906	nuclear phosphop
19	80.5	8.1	1076	D82083	carbamoyl-phosphat
20	80	8.0	851	AG0482	probable hybrid tw
21	79.5	8.0	461	H69350	glycolate oxidase
22	79.5	8.0	859	F85680	unknown protein en
23	79.5	8.0	859	C90853	probable portal pr
24	79.5	8.0	859	H90909	ABC transporter ho
25	79	7.9	316	C69855	hypothetical prote
26	79	7.9	992	T08772	transcription regu
27	78.5	7.9	344	C72395	serine/threonine k
28	78.5	7.9	360	S49330	conserved hypotet
29	78.5	7.9	394	B72419	

ALIGNMENTS

RESULT 1

T03476

conserved hypothetical protein - Rhodobacter capsulatus

C;Species: Rhodobacter capsulatus

C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 08-Oct-1999

C;Accession: T03476

R;Vicek, C.; Paces, V.; Maltsev, N.; Paces, J.; Haselkorn, R.; Fonstein, M.

Proc. Natl. Acad. Sci. U.S.A. 94, 9384-9388, 1997

A;Title: Sequence of a 189-kb segment of the chromosome of Rhodobacter capsulatus SB1003

A;Reference number: Z14955; MUID:97404404; PMID:9256491

A;Accession: T03476

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-608 <VLC>

A;Cross-references: EMBL:AF010496; NID:g3128256; PIDN:AAC16129.1; PID:g3128277

C;Genetics:

A;Map position: 1

Query Match 10.2%; Score 102; DB 2; Length 608;
Best Local Similarity 37.3%; Pred. No. 1.1;
Matches 28; Conservative 10; Mismatches 29; Indels 8; Gaps 3;

QY	86	GWKVIFF----	KTPNQDTEFLRLNLFLEKEGQTVSGMFRALGQRLSPATVPCISPELL 141
DB	407	GYWRYFHDWIDTPNPEAQMILASVMFDLRPIGGTGADLFTALQERTLRLAAQ--SPIFI 463	
QY	142	AHLGQAMAHAPQPL 156	
DB	464	AHMLGNALKHTP-PL 477	

RESULT 2

C72403

hypothetical protein - Thermotoga maritima (strain MSB8)

C;Species: Thermotoga maritima

C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000

C;Accession: C72403

R;Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey, Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;

C.M.

Nature 399, 323-329, 1999

A;Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq

A;Reference number: A72200; MUID:99287316; PMID:10360571

A;Accession: C72403

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-341 <ARN>

A;Cross-references: GB:AE001707; GB:AE000512; NID:g4980720; PIDN:AA035332.1; PID:g498073

A;Experimental source: strain MSB8

C;Genetics:

A;Gene: TM0241

capsular polysacch
E1 protein - human
hypothetical prote
protein C14H10.2
probable transport
hypothetical prote
glucosylceramidase
conserved hypotet
aminotripeptidase
triglyceridase
trigger factor lim
urban - silkworm
hypothetical prote
FtsZ protein - Myc
NADH dehydrogenase

30	78.5	7.9	420	2	G89777
31	78.5	7.9	606	1	M1WL5
32	78.5	7.9	607	2	T01904
33	78.5	7.9	680	2	D89625
34	78.5	7.9	745	2	G02500
35	78.5	7.9	746	2	T19287
36	78.5	7.9	831	2	JC7880
37	78	7.8	117	2	A83312
38	78	7.8	273	2	C83447
39	78	7.8	409	2	A42363
40	78	7.8	409	2	AD0646
41	78	7.8	471	2	AB2366
42	78	7.8	551	2	S52487
43	78	7.8	1418	2	S40764
44	77.5	7.8	380	2	JC7835
45	77.5	7.8	409	2	AG2317

A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: F83080
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-751 <STO>
A;Cross-references: GB:AE004866; GB:AE004091; NID:g9950760; PIDN:AAG07911.1; GSPDB:GN001
A;Experimental source: strain PA01
C;Genetics:
A;Gene: PA4523

Query Match 8.5%; Score 85; DB 2; Length 751;
Best Local Similarity 23.3%; Pred. No. 38;
Matches 50; Conservative 35; Mismatches 76; Indels 54; Gaps 11;

QY 4 LLEDWCRIMSVDEOKSLMTGIPADFEAEIQ-EVLQETLKSIGRYLLGKIFRKQENAN 62
DB 449 LLEDFTAFGTGDERRRSELLEQRTDRAEGRKAELARQDVEQVLNQLLKGVLPE----- 503
QY 63 AVILLELLED--TDVSAIPSEVQKGGV-MKVIFKT-----PQDTEFLERLNLFL 110
DB 504 -VVRLLRDAWSKVMLLICLKHGKASAEWREALDTWDELIVSWGPHDDP--LSRQQL-- 559
QY 111 KEGQTVSGMFRALGQERALSPTVPCISPELLA-----HLG-----QAMAHAP----- 153
DB 559 ---ERVPGLLKALREGLASAFDPFSTGTFPSQLETLHVQSLQDKRGEAAVADSPLLDE 615
QY 154 -----QPLPMRYR---KLRVFGSVAVPAPPEE 177
DB 616 ASLGTAAEKPRAPVWLVEVEIVLTSPDAAPAVE 650

RESULT 5
140486
Surfactin synthetase component II - Bacillus subtilis
N;Alternate names: surfactin synthetase srfA2; surfactin synthetase/competence protein
N;Contains: acid-amino-acid ligase (EC 6.3.2.-)
C;Species: Bacillus subtilis
C;Date: 12-Aug-1996 #sequence revision 12-Aug-1996 #text change 03-Nov-2000
C;Accession: 140486; S60866; C69718; S46968; S35518; S25658; S34986
R;Cosmina, P.; Rodriguez, F.; de Ferra, F.; Grandi, G.; Perego, M.; Venema, G.; van Sin
Mol. Microbiol. 8, 821-831, 1993
A;Title: Sequence and analysis of the genetic locus responsible for surfactin synthesis
A;Reference number: 140485; MUID:93360813; PMID:8355603
A;Accession: 140486
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-3587 <RES>
A;Cross-references: EMBL:X70356; NID:g396480; PIDN:CAA49817.1; PID:g396482
A;Experimental source: strain W168 derivative of JH642
R;Hamoen, L.W.; Eshuis, H.; Jongbloed, J.; Venema, G.; van Sinderen, D.
Mol. Microbiol. 15, 55-63, 1995
A;Title: A small gene, designated comS, located within the coding region of the fourth
A;Reference number: S60866; MUID:95272393; PMID:7752896
A;Accession: S60866
A;Molecule type: DNA
A;Residues: 977-1104 <HAM>
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E
Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, V.; Fuma, S.; Galizzi, A.; Galie
Riegler, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Ser
akeuchi, M.; Tanakoshi, A.; Tanaka, T.; Terpsira, P.; Tognoni, A.; Tosato, V.; Uchiyama
T.; Winters, P.; Wipat, A.; Yamamoto, E.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida,
A;Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis
A;Reference number: A69580; MUID:98044033; PMID:9384377
A;Accession: C69718

Query Match 9.5%; Score 95; DB 2; Length 341;
Best Local Similarity 24.1%; Pred. No. 2.1;
Matches 48; Conservative 43; Mismatches 72; Indels 36; Gaps 12;

QY 4 LLEDWCRIMSVDEOKSLMTGIPADFEAEIQ-EVLQETLKSIGRYLLGKIFRKQENANA 63
DB 121 LFEENSEKLEVE-----VGVPLDYEDVSL-EVVRKEHFEELSTDEVF-KLQREDLRLN 172
QY 64 ----VLELLEDTDVSAIPSEVQKGGVVKVFKTPNQDTEFLERLNLFL-EKEG-OTVS 117
DB 173 PEVFPVLDPLNGVTIREIDVEER-----IDPVILNYGNE--EPSEKPKFLKFKGTENVA 226
QY 118 GM-FRALGQERALS--PATVPCISPELLAHLGQAMAHAPOLPLPMRYKLRVFGSVA 174
DB 227 SLPGKVLSEIKVTRGTYLLKVEFEHIGIAIV-----SPNLKIKRGSERV----- 275
QY 175 PEESEFVWLEQATEIVKE 193
DB 276 --KISDEEMKEIGEMLKE 292

RESULT 3
T02632
hypothetical protein DJ1186C01.1 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 05-Mar-1999 #sequence revision 05-Mar-1999 #text_change 05-Nov-1999
C;Accession: T02632
R;Bemis, G.; Rohlfing, T.; Morris, M.
submitted to the EMBL Data Library, July 1998
A;Description: The sequence of Homo sapiens PAC clone DJ1186C01.
A;Reference number: Z14682
A;Accession: T02632
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-1684 <BEM>
A;Cross-references: EMBL:AC004991; NID:g3342751; PIDN:AAC27675.1; PID:g3342752
C;Map position: 7
A;Note: intron positions not resolved (incomplete sequence)
A;Note: WUGSC:H_DJ1186C01.1

Query Match 8.8%; Score 87.5; DB 2; Length 1684;
Best Local Similarity 24.0%; Pred. No. 63;
Matches 50; Conservative 28; Mismatches 71; Indels 59; Gaps 9;

QY 20 LMVTGIPADFEAEIQEVLQETLKSIGR-----YRLGKIFRKQENANAVLL 66
DB 684 LCVNQLQDFLYDHIQPVRAELMQALWRLTRNPADSIHVAYRVLGKF---GGSNRKML 739
QY 67 ELLEDTDVSAIPSEVQKGGVVKVFKTPNQDTEFL---RLNLFLEKEGQT----- 115
DB 740 K-ESQKLVYVTEVQ-----PSITVEFSDCKASQLPMEKAIETALDCLKSA 786
QY 116 -VSGMFRALGQERALSPTVPCISPELLAHLGQAMAH-----APQPLPMRYKLRV 167
DB 787 NTEPYRQAEVVKCFVAMSLSDNKHLYQLLAHFNFTKTPNVIISHRYK----- 841
QY 168 SGSVAPEPEESFEVWLEQA--TEIVKE 193
DB 842 ---AQDTPARKTFEQALGATGAPMSAVIKD 866

RESULT 4
F83080
hypothetical protein PA4523 [imported] - Pseudomonas aeruginosa (strain PA01)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence revision 15-Sep-2000 #text_change 31-Dec-2000
C;Accession: F83080
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warriner, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho

A>Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-3587 <KUN>

A:CROSS-references: GB:299105; GB:AL009126; NID:G2632457; PIDN:CAB12143.1; PID:G2632635
 A:Experimental sources: Strain 168
 R:Fabret, C.; Quentin, Y.; Guiseppi, A.; Busuttill, J.; Halech, J.; Denizot, F.
 Submitted to the EMBL Data Library, March 1993
 A:Reference number: S46967
 A:Accession: S46968

A:Molecule type: DNA
 A:Residues: 1-32 'F', 34-41 'G', 43-109 'D', 111-114 'G', 116-138 'V', 140-258 'W', 260-308 'A', 1756-1914 'PK', 1917-2138 'SRLN', 2146-2244 'Q', 2446-2712 'H', 2714-2722 'H', 272
 A:CROSS-references: EMBL:X72672; NID:G516358; PIDN:CAA51223.1; PID:G516360
 R:Fuma, S.; Fujishima, Y.; Corbelli, N.; D'Souza, C.; Nakano, M.M.; Zuber, P.; Yamane, K.
 Nucleic Acids Res. 21, 93-97, 1993

A>Title: Nucleotide sequence of 5' portion of srfA that contains the region required for

A:Reference number: S35517; MUID:93181186; PMID:8441623

A:Accession: S35518

A>Status: significant sequence differences

A:Molecule type: DNA

A:CROSS-references: EMBL:DL3262; NID:G216345; PID:G216347

A:Experimental sources: Strain 168 trpC2

R:Borchert, S.; Patil, S.S.; Marahiel, M.A.
 FEMS Microbiol. Lett. 92, 175-180, 1992

A>Title: Identification of putative multifunctional peptide synthetase genes using high

A:Reference number: S25658

A:Accession: S25658

A:Molecule type: DNA

A:Residues: 168 'C', 170-171, 281-283, 514-595, 597-647 'R', 649-679 'ETL', 683-693 'DKR', 697,

A:CROSS-references: EMBL:X65835; NID:G40202; PIDN:CAA46678.1; PID:G40203

A:Experimental sources: Strain ATCC 21332

C:Comment: This protein contains several amino acid-activating domains for the synthesis

the amino-terminal region of this protein, appear to be required for the development of

C:Genetics:

A:Gene: srfAB, srfA2

C:Superfamily: surfactin synthetase; acetate-CoA ligase homology; acyl carrier protein h

C:Keywords: antibiotic biosynthesis; carrier protein; duplication; ligase; phosphopantet

F:511-951/Domain: acetate-CoA ligase homology <AC11>

F:968-1035/Domain: acyl carrier protein homology <ACPI>

F:1036-1481/Domain: repeat <RPT1>

F:1542-1995/Domain: acetate-CoA ligase homology <ACI2>

F:2013-2081/Domain: acyl carrier protein homology <ACP2>

F:2082-2529/Domain: repeat <RPT2>

F:2591-3024/Domain: acetate-CoA ligase homology <ACI3>

F:3041-3108/Domain: acyl carrier protein homology <ACP3>

F:999,2045,3073/Binding site: phosphopantetheine (Ser) (covalent) #status predicted

Query Match 8.5%; Score 84.5; DB 2; Length 3587;

Best Local Similarity 18.9%; Pred. No. 2.9e+02;

Matches 44; Conservative 44; Mismatches 82; Indels 63; Gaps 8;

QY 11 LMSVDEQKSLMTGIPADFEAEIQVLOETLKSGLRYLLGKIFRKQENANAVILLE 70

Db LSKISGQDIVGSVTAGRTNADVDQPMGFNTIAL-----RMEAKEQQTFAELLELAK 2399

QY 71 DTVSAIPSEV-----QKGGVWKVIFKTPNODTFLERLNI----- 107

Db QTNLSALEHQEYPPFDLVNQLDPRDMSRNFNVWVTENPDKEILTQNLISISPYRAH 2459

QY 108 -----FLKEG-----QTVSGMFRALGOEALSPATVPCISPELLAHLGQAMA 150

Db QGTSKFDLTGLGFTDENGIGLQLEVATDLFAKETAKWS-----EYVLRLL-KAVA 2509

QY 151 HAP-QPLLP-----RYRKLVPSSGSAVPAPEESFEVWLEQATEIVKWP 195

Db 2510 DNPQNQLSLLVLTETEQALLKAWKGAIPVPTDKTVHQLFEETVQRHKORP 2562

RESULT 6

I64034

hypothetical protein H11522 - Haemophilus influenzae (strain Rd Kw20)

C:Species: Haemophilus influenzae

C:Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 08-Oct-1999
 A:Accession: I64034

R:Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, J.
 ; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, C.
 ; D.M.; Brandon, R.C.; Pine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.
 Science 269, 496-512, 1995

A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
 A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.

A:Reference number: A64000; MUID:95350630; PMID:7542800

A:Accession: I64034

A>Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-623 <ITGR>

A:CROSS-references: GB:U32828; GB:I42023; NID:G1574362; PIDN:AAAC23178.1; PID:G1574363;

Query Match 8.4%; Score 83.5; DB 2; Length 623;

Best Local Similarity 21.7%; Pred. No. 40;

Matches 33; Conservative 19; Mismatches 51; Indels 49; Gaps 6;

QY 7 DWCRIMSVDEQKSLMTGTIPAD-----FEABIQVLOETLKSGLRYLLG- 52

Db 41 DMFNMVQAEFLFSILDAGIQPKGRNLNQISLAIRKLSEKGVDFSOQ-LKQADGKYVIGR 99

QY 53 -----KIFRKQENANAVLL-----ELLEDDTVSAIPSE-----VQGGGV 87

Db 100 CKSVAEALRTIRPTENGQRILVDAYEGTAGGGEFVADLQDLITPDGGTCFVVPNNNGR 159

QY 88 WKVIFKTPNQDTEF-----LERLNLFLE 110

Db 160 WKRLFSLSQDTEFGVIGGVADDTTNLNAFLD 191

RESULT 7

phytochrome B - common tobacco

C:Species: Nicotiana tabacum (common tobacco)

C:Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 21-Jan-2000

A:Accession: T03668; T03672

R:Kern, R.; Gasch, A.; Deak, M.; Kay, S.A.; Chua, N.H.

Plant Physiol. 102, 1363-1364, 1993

A>Title: The cucumber long hypocotyl mutant lacks a light-stable PHYB-like phytochrome.

A:Reference number: Z14996; MUID:94105358; PMID:8278560

A:Accession: T03668

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1132 <KER>

A:CROSS-references: EMBL:L10114; NID:G295345; PIDN:AAA34092.1; PID:G295346

A:Experimental sources: strain SR1; tissue-type etiolated seedling

R:Lopez-Juez, E.; Nagatani, A.; Tomizawa, K.; Deak, M.; Kern, R.; Kendrick, R.E.; Furuya

Plant Cell 4, 241-251, 1992

A>Title: The cucumber long hypocotyl mutant lacks a light-stable PHYB-like phytochrome.

A:Reference number: Z14997; MUID:92361250; PMID:1498594

A:Accession: T03672

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 457-506 'S', 508-586 'Q', 587-1132 <LOP>

A:CROSS-references: EMBL:M65023; NID:G170286; PIDN:AAA34093.1; PID:G170287

C:Genetics:

A:Gene: phyB

C:Superfamily: phytochrome; phytochrome homology

C:Keywords: chromoprotein; photoreceptor; phytochromobilin

F:80-589/Domain: phytochrome homology <PHY>

F:336/Binding site: phytochromobilin (Cys) (covalent) #status predicted

Query Match 8.3%; Score 83; DB 2; Length 1132;

Best Local Similarity 21.6%; Pred. No. 92;

Matches 43; Conservative 28; Mismatches 72; Indels 56; Gaps 8;

QY 2 LALLEDCRIMSVDEQKSLMTGIPADFEAEIQVLOETLKSGLRYLLGKIFRKQENA 61

Db 586 LLILRDSFKDAEASNSKAV----VHAQLGENELQGI--DELSSVAR-----EMVRLIETA 634

QY 62 NAVLL-----ELLEDDTVSAIPSEVQGGGVWKVIFKTPNQDTEFLERLNL 108

Db 635 TAPFAVDVEGRINGWNAKVAETDLVS--EEAMGKSLVDVHVKESQETAELK-LFNAL 691

QY 109 LEKEGOTVSGMFRALQGEALSPATV----PCISPELLAHLGLG----- 146

Db 692 RGEEDKNVEIKLRTGPEQLKKAFFVWVWACSSKDYTNVGVFCVGDQVTGQKVVMDKF 751

QY 147 -----QAMAHAPQPLLP 158

Db 752 IHIQGDYKAIVHSPNPLIP 770

RESULT 8

T21720

hypothetical protein F34D10.6 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T21720

R:Kershaw, J.

submitted to the EMBL Data Library, June 1994

A:Reference number: Z19464

A:Accession: T21720

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1295 <WIL>

A:Cross-references: EMBL:Z34799; PIDN:CAA84316.1; GSPDB:GN00021; CESP:F34D10.6

A:Experimental source: Clone F34D10

C:Genetics:

A:Gene: CESP:F34D10.6

A:Map position: 3

A:Introns: 7/3; 46/3; 109/1; 204/2; 227/3; 266/3; 298/2; 336/1; 397/1; 429/3; 512/2; 597

Query Match 8.3%; Score 83; DB 2; Length 1295;

Best Local Similarity 24.9%; Pred. No. 1.1e+02;

Matches 49; Conservative 26; Mismatches 60; Indels 62; Gaps 11;

QY 35 QEVLTQELKSGRYRLA-----GKIFRKQENANAVLLELDTDVSAIPS 79

Db 442 QNAIYSLVDLIGARLDADQGITAEYRQNPGRLLRPSSAVSSVWLRSMTDDEFFDPG 501

QY 80 EVQKGKGVW-----KVIFKTPNQDTFELRLNLFLEK---EGQTVSGMF 120

Db 502 E-QSQGAGMSKQFLSTFSVVPNHKVIKFK---KNKFFLNLTITPFYKILLEYVTLDFV 557

QY 121 RALQGEALSPATVPCISPELLAHLGLQAMAHAPQPLLPMPYRKLRVFGSSAV----PAPE 176

Db 558 BAAGRDAY-----YLAGISSVHI-KIMLNESEFKKL--FSPFFILNPPNAPS 600

QY 177 EESFEVMLEQAT-EIVK 192

Db 601 EEQVEEWL--ATGEVLK 615

RESULT 9

T06694

hypothetical protein T17F15.240 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 22-Oct-1999

C:Accession: T06694

R:Quetier, F.; Choisme, N.; Robert, C.; Brottier, P.; Wincker, P.; Cattolico, L.; Artigou

submitted to the Protein Sequence Database, April 1999

A:Reference number: Z15793

A:Accession: T06694

A:Molecule type: DNA

A:Residues: 1-1528 <QUE>

A:Cross-references: EMBL:AL049658; GSPDB:GN00661; ATSP:T17F15.240

A:Experimental source: cultivar Columbia; BAC clone T17F15

C:Genetics:

A:Gene: ATSP-T17F15.240

A:Map position: 3

A:Introns: 823/3; 866/3; 930/3; 947/2; 1051/3; 1094/1; 1161/1; 1193/1; 1253/3; 1325/3; 1

Query Match 8.3%; Score 83; DB 2; Length 1528;

Best Local Similarity 27.4%; Pred. No. 1.3e+02;

Matches 49; Conservative 26; Mismatches 72; Indels 32; Gaps 10;

QY 28 DPEEABIQEVLQET-LKSLGRYRLGLKIFRKQENANAVLLELDTD---VSAIPSEVQG 83

Db 1174 DDEERFQADLKRKALQSL-----DVGRRNMTSLCKTSLDNDNWLSDVTKESQS 1226

QY 84 KGGWVKVIFKTPNQDTFELRLNLFLEKGGQTV--SGMFRALQGEALSPATV-----P 134

Db 1227 SPGV--AIFGTGLQNE--VGEYCNFLNVIQSLWNLGMFRA---EFLRSSTLHHHFGDP 1279

QY 135 CISPPELLAHLGLQAMAHAPQPLLPMPYRKLRVFGSSAVPAPEESP--EVLWQATEIV 191

Db 1280 CVVCSLYAIFTALSTASSTRKEPVPASSLRIALSNLYP---DSSFFQEQAMNDASEVL 1335

RESULT 10

E69823

conserved hypothetical protein yhcQ - Bacillus subtilis

C:Species: Bacillus subtilis

C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 15-Oct-1999

C:Accession: E69823

R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte

C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch

A.; Ehrlich, S.D.; Emmerson, P.F.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.

Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritz, C.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallen

iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holstappel, S.; Hosono, S.; Hulio, M.F

Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois

A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mauee

Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetell

Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon

A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Sero

akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama

T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida,

A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.

A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.

A:Reference number: A69580; MUID:98044033; PMID:9384377

A:Accession: E69823

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-217 <KUN>

A:Cross-references: GB:Z99108; GB:AL009126; NID:g2633055; PIDN:CAB12746.1; PID:el182907

A:Experimental source: strain 168

C:Genetics:

A:Gene: yhcQ

Query Match 8.2%; Score 81.5; DB 2; Length 217;

Best Local Similarity 25.0%; Pred. No. 16;

Matches 46; Conservative 33; Mismatches 76; Indels 29; Gaps 8;

QY 16 EQKSLMVTGIP-----ADFEABIQEVLQETLKSGRYRLGLKIFRKQENANAVLLEL 68

Db 7 QQSQSMKGIKPKHKNHGHEMFMHEVLSLTFLVDQFMWLRQFCKDQE-----LLNI 61

QY 69 LEDTDVSAIPSEVQKGGWVKVFKT---PNQDTFELRLNLFLEKGGQTVSGMFRALQ 125

Db 62 L-DROHQFITSQY-----NITACEFKTGPSEPKT-----ATYMKEDQTVYGMQPSQPK 111

QY 126 ---EALSPATVPCISPELLAHLGLQAMAHAPQPLLPMPYRKLRVFGSSAVPAPEESP 182

Db 112 KPVQSMNDIDDSIISRQMLCAIKAQA-SMLTWASLEMTNPVRRVLSAQIQEYVEMAFEI 170

QY 183 MLEQ 186

Db 171 FLYQ 174

RESULT 11

F69188

ethylene-inducible protein - Methanobacterium thermoautotrophicum (strain Delta H)

C:Species: Methanobacterium thermoautotrophicum

C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 17-Nov-2000

C;Accession: F69188
F;Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiawani, N;
ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A;Title: Complete genome sequence of *Methanobacterium thermoautotrophicum* Delta H: funct
A;Reference number: A69000; MUID:98037514; PMID:9371463
A;Accession: F69188
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-293 <MTH>
A;Cross-references: GB:AE000846; GB:AE000666; NID:g2621740; PIDN:ARB85171.1; PID:g262174
A;Experimental source: strain Delta H
C;Genetics:
A;Gene: MTH666
C;Superfamily: hypothetical protein HI1647

Query Match	8.2%;	Score 81.5;	DB 2;	Length 293;
Best Local Similarity	21.7%;	Pred. No. 23;		
Matches	39;	Conservative 30;	Mismatches 50;	Indels 61; Gaps 9
QY	59	ENANAVLLELLEDTDSAIPSEVGCGKGVWVIKTNPQDTEFL-----RLNLF	108	
Dd	34	EDSGAVMALE-----KVPADIRASGGVARM--ADPNKVQEIMDAIVSPVMKVRIGHF	86	
QY	109	LXEGGTVSGMFRALG-----QBALSPA-----TPVCISPELLAHLIGQAMAH	151	
Dd	87	VEAQ-----VLEALGVDMIDEVLTPADERFHIDKKFTVPFCV-----ARNIGEALRR	137	
QY	152	APQPLPMR-----YRKLVFGSVAIPAEEBSFEVW-----LEQATEIVKE	193	
Dd	138	IDEGAAWRTKGPGTCGNIVEAVRHMKIMMSREIQNKCEEELWEVSRIEAPLELVRE	197	

RESULT 12

S18609 glutamate dehydrogenase (NADP) (EC 1.4.1.4) - Halobacterium salinarum
N1:Alternate names: glutamic dehydrogenase; NADP-specific glutamate dehydrogenase
C1:Species: Halobacterium salinarum
C1:Date: 13-Jan-1995 #sequence_revision 23-Mar-1995 #text_change 03-Jun-2002
C1:Accession: S18609
C1:Accession: R.R.Benachennou, N.; Baldacci, G.
Mol. Gen. Genet. 230, 345-352, 1991
A1:Title: The gene for a halophilic glutamate dehydrogenase: sequence, transcription anal
A1:Reference number: S18609; MUID:92114863; PMID:1766432
A1:Accession: S18609
A1:Molecule type: DNA
A1:Residues: 1-435 <BENS>
A1:Cross-references: EMBL:X63837; NID:G49045; PIDN:CAA45327.1; PID:G49046
A1:Note: the source is designated as Halobacterium salinarum
C1:Superfamily: glutamate dehydrogenase (NAD(P)+)
C1:Keywords: NADP; oxidoreductase
F1:126/Binding site: substrate (Lys) #status predicted

```
Query Match      8.2%; Score 81.5; DB 1; Length 435;
Best Local Similarity 25.6%; Pred. No. 37;
Matches 41; Conservative 19; Mismatches 59; Indels 41; Gaps 7;
```

2y 4 LLEDW-CRIMSVDEQKSLM-----VTGIPADFEEAEIQVLTLSLGRY-----48
||| : | : : : : : : : : : : : : : : : :
bb 249 LLDKNGATIVAIISDVGAMYEPDGIDTASVPSHDEEPAITYADTVISNEELLTLDVDV 308

2y 49 ---RLICKIRKCKENANAVLLEI-----LETDYSALPSEVOQKGGMVK 89
||| : | : : : : : : : : : : : : : : : :
bb 309 LI PAALGNVITK-ENAEATAADLVVEGANGPTTADSILADRDAVPIPDILANAGGVTV 367

2y 90 VIF---KTPNQDTFFELRNLFLEKEGQTVSGMFALGOE 126
||| : | : : : : : : : : : : : : : : : :
bb 368 SYFEWLQDINRAWSLRVNDELEAEMQAA---WRAVKDE 404

RESULT 13
G90256

DNA primase, probable [imported] - *Sulfolobus solfataricus*
C:Species: *Sulfolobus solfataricus*
C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001
C:Accession: G90256
R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder,
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A:Description: *Sulfolobus solfataricus* complete genome.
A:Reference number: A99139
A:Accession: G90256
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-330 <KUR>
A:Cross-references: GB:AE006641; NID:gl3814232; PIDN:AAK41310.1; GSPDB:GN00155
C:Genetics:
A:Gene: SSO1048

Query Match	8.1%;	Score 81;	DB 2;	Length 330;
Best Local Similarity	26.1%;	Pred. No. 29;	Mismatches	31; Conservative 15;
				Gaps 6;
QY	9	CRIMSVDEQKSLM-----VTGIPADFEAEITQEVQLQETLKSILGRYLRLKIFRKOENA	61	
	:: :: :	:: :: :	:: :: :	:: :: :
Dd	188	CALLDSDERKEIAEYVMGIGVPGCGSENAP-----GWGRKNRG	229	
QY	62	NAVLLLELLEDTVS---AIPSEVOGGGVWVKVIKFTPN-ODTFELERLNLFLEKGQTV	116	
	:: :	:: :: :	:: :: :	:: :: :
Dd	230	NGVTIDQVTDIKRLIRPNSLHGKSL--IVKRVPLNDDFNETLSFP--TGYTI	283	

RESULT 14
AE2576

hypothetical protein alr9504 [imported] - Nostoc sp. (strain PCC 7120) plasmid pCC7120se
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C:Accession: AE2576
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,
Nakazaki, N.; Shimo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A:Reference number: AB1807; UID:21595285; PMID:11759840

A:Cross-references: SB:AP003606, PIDN:BAW7521.1; PID:gl7134967; GSEDB:GN00185
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: alr9504
A:Genome: plasmid

Query Match		8.1%;	Score 81;	DB 2;	Length 390;
Best Local Similarity		22.5%;	Pred. No. 36;		
Matches		Conservative 36;	Mismatches 74;	Indels 66;	Gaps 12;
QY	15	DEQSLMVTGIPADPEE-----ABIQE-----	VLOETUKGLG	46	
dDb	158	DDWQGLKVVGKGDGEINASNNADOIDLENRIKURSQSGDNALTKELCIITAEFBFPSLK	:: :: :: :: ::	217	
QY	47	-----RYRLLGKIFPKOENANAVLLELLEDTDVSAIPEVSQCKGGVKVIKFTPNODTEF	:: :: :: :: ::	101	
dDb	218	DECEIAPEWLGLKLARRGRKPIFVIALSQSDSVKAL--GIEGDGI-----RQNFKY	:: :: :: :: ::	267	
QY	102	LERNLNFLEKGQTV-----SCMFRALQOEALSPATVPCTISPELLAH---LLGOA	:: :: :: :: ::	148	
dDb	268	L-RLGKFVAHQAKKLGNDAVMNWNLSQGYIRAWLDD--EPQAQMPSDSTLLVHTRLLDSRA	:: :: :: :: ::	324	
QY	149	MA--HAQOPLLPMRYKRLRVTFSSAVPAPEESFEVFWLEQATEIVKE	:: :: :: :: ::	193	
dB	325	IAGNPPEPPPEPMDLTNLPHSLKPAILALKOBG---WSD--SKLIKE	:: :: :: :: ::	366	

Search completed: September 21, 2004, 13:40:39
Job time : 16.7398 secs

GenCore version 5.1.6

Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 13:20:21 ; Search time 5.31599 Seconds
(without alignments)
1910.028 Million cell updates/sec

Title: US-10-037-860-7

Perfect score: 996

Sequence: 1 P1ALLEDCRMSVDQKSL.....EESFEVWLFQATEIVKEWP 195

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	982	98.6	364	1	PM22_HUMAN	Q9u142 homo sapien
2	959	96.3	364	1	PM22_MACFA	Q9gm03 macaca fasc
3	766	76.9	365	1	PM22_MOUSE	Q8bhk0 mus musculus
4	468	47.0	353	1	PM21_RAT	Q8vzh4 rattus norv
5	466	46.8	353	1	PM21_HUMAN	Q8nd90 mus musculus
6	459	46.1	353	1	PM21_MOUSE	Q8clc8 mus musculus
7	436	43.8	351	1	MOP1_HUMAN	Q96by2 homo sapien
8	431	43.3	351	1	MOP1_MACFA	Q95ki4 macaca fasc
9	431	43.3	352	1	MOP1_MOUSE	Q9erh6 mus musculus
10	91	9.1	2298	1	CU05_HUMAN	Q9v1r5 homo sapien
11	87.5	8.8	3859	1	TRAP_HUMAN	Q9v4a5 homo sapien
12	84.5	8.5	520	1	CET1_CANAL	Q91803 candida alb
13	84.5	8.5	3587	1	SFP2_BACSU	Q04747 bacillus su
14	83.5	8.4	623	1	VPS_HAEIN	P42424 haemophilus
15	83	8.3	1132	1	PHYB_TOBAC	P29130 nicotiana t
16	82.5	8.3	886	1	PHYB_MOUSE	Q8t5l3 mus musculus
17	82	8.2	452	1	IM44_HUMAN	O43615 homo sapien
18	81.5	8.2	217	1	YHQQ_BACSU	P4601 bacillus su
19	81.5	8.2	293	1	Y666_METTH	Q26762 methanobact
20	81.5	8.2	435	1	DH4_HALSA	P29051 halobacteri
21	81.5	8.2	820	1	BCD2_MOUSE	Q921c5 mus musculus
22	81.5	8.2	1077	1	CARB_VIBPA	Q87sf3 vibrio para
23	81.5	8.2	1077	1	CARB_VIBVU	Q8dem2 vibrio vuln
24	81	8.1	330	1	PRIS_SULSO	Q97z83 sulfolobus
25	81	8.1	502	1	C3A_ARATH	P48421 arabidopsis
26	81	8.1	555	1	ODP2_RAT	P08461 rattus norv
27	80.5	8.1	489	1	RNF9_MOUSE	Q9wh5 mus musculus
28	80.5	8.1	559	1	DNLI_PYRKO	Q9hhc4 pyrococcus
29	80.5	8.1	1076	1	CARB_VIBCH	Q9kph9 vibrio chol
30	80	8.0	421	1	FXJ1_HUMAN	Q92949 homo sapien
31	80	8.0	1130	1	PHYB_SOLTU	P34094 solanum tub
32	79	7.9	1016	1	EM11_HUMAN	Q9y6c2 homo sapien
33	78.5	7.9	360	1	CDKA_HUMAN	Q15131 homo sapien

Q13043 homo sapien
P06920 human papil
P33897 homo sapien
Q913c3 pseudomonas
Q827h6 salmonella
P26311 salmonella
Q8yq99 anabaena sp
P34641 caenorhabdi
P26542 human papil
P49035 daucus caro
P46940 homo sapien
P21263 rattus norv

ALIGNMENTS

RESULT 1
PMA2_HUMAN
ID PMA2_HUMAN STANDARD; PRT; 364 AA.
AC Q9UL42; O94959; O95145; Q9UL43;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Paraneoplastic antigen M2 (Onconeural antigen MA2) (Paraneoplastic
DE neuronal antigen MM2) (40 kDa neuronal protein).
GN PMA2 OR MA2 OR KIAA0883.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RP TISSUE=Cerebellum;
RX MEDLINE=99270611; PubMed=10362822;
RA Voltz R., Gultekin S.H., Rosenfeld M.R., Gerstner E., Eichen J.,
RA Posner J.B., Balmou J.,
RT "A serologic marker of paraneoplastic limbic and brain-stem
encephalitis in patients with testicular cancer.";
RL New Engl. J. Med. 340:1788-1795(1999).
[2]
SEQUENCE FROM N.A.
RP TISSUE=Brain;
RX MEDLINE=99156230; PubMed=10048485;
RA Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirose M.,
RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XII.
The complete sequences of 100 new cDNA clones from brain which code
for large proteins in vitro.";
RL DNA Res. 5:355-364(1998).
[3]
IDENTIFICATION, AND SUBCELLULAR LOCATION.
RX MEDLINE=99158179; PubMed=10050892;
RA Dalmay J., Gultekin S.H., Voltz R., Hoard R., DesChamps T.,
RA Balmaceda C., Bachelot T., Gerstner E., Eichen J., Frennier J.,
RA Posner J.B., Rosenfeld M.R.;
RT "Mal, a novel neuron- and testis-specific protein, is recognized by
the serum of patients with paraneoplastic neurological disorders.";
RL Brain 122:27-39(1999).
CC -!- SUBCELLULAR LOCATION: Nuclear; nucleolar.
CC -!- TISSUE SPECIFICITY: Brain specific. In some patients suffering
from cancers, it is also specifically expressed by the testicular
tumor cells.
CC -!- MISCELLANEOUS: Antibodies against PMA2 are present in sera from
patients suffering of paraneoplastic neurological disorders.
CC -!- SIMILARITY: Belongs to the PMA family.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)

or send an email to license@isb-sib.ch.

CC EMBL; AF037365; AAD02098.1; -
 CC EMBL; AF083114; AAF05625.1; -
 DR EMBL; AF083115; AAF05626.1; -
 DR EMBL; AF286487; AAG28165.1; -
 DR EMBL; AB020690; BAA74906.1; -
 DR Genew; HGNC:9159; PNMA2.
 DR MIM; 603970; -
 KW Antigen; Tumor antigen; Nuclear protein.
 FT DOMAIN 333 338 POLY-GLU.
 FT CONFLICT 128 129 GV -> AL (IN REF. 1; AAD02098).
 FT CONFLICT 141 141 L -> I (IN REF. 1; AAD02098).
 FT CONFLICT 257 257 T -> P (IN REF. 1; AAF05625).
 FT CONFLICT 278 278 R -> K (IN REF. 1; AAF05626).
 FT CONFLICT 364 364 6E417AD96E3F0E93 CRC64;
 SQ SEQUENCE 364 AA; 41509 MW; 6E417AD96E3F0E93 CRC64;
 Query Match 98.6%; Score 982; DB 1; Length 364;
 Best Local Similarity 99.0%; Pred. No. 2.8e-76;
 Matches 192; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 2 LALLEDCWCRMSVDEQKSLMVTGIPADFEAEIQEVLQETLKSIGRYRLGKIFRKQENA 61
 DB 3 LALLEDCWCRMSVDEQKSLMVTGIPADFEAEIQEVLQETLKSIGRYRLGKIFRKQENA 62
 QY 62 NAVLELLEDDTDSAIPTSEVQGGVWKVIFKTPNQDTFELRLNLFLEKGTQVSGMFR 121
 DB 63 NAVLELLEDDTDSAIPTSEVQGGVWKVIFKTPNQDTFELRLNLFLEKGTQVSGMFR 122
 QY 122 ALGOEALSPATVPCISPELLAHLLGQMAHAPQLPMPYRKLRFVSGSAVPAPEESFE 181
 DB 123 ALGOEGVSPATVPCISPELLAHLLGQMAHAPQLPMPYRKLRFVSGSAVPAPEESFE 182
 QY 182 VWLEQATEIVKEWP 195
 DB 183 VWLEQATEIVKEWP 196
 Query Match 99.0%; Score 982; DB 1; Length 364;
 Best Local Similarity 99.0%; Pred. No. 2.8e-76;
 Matches 192; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 2 LALLEDCWCRMSVDEQKSLMVTGIPADFEAEIQEVLQETLKSIGRYRLGKIFRKQENA 61
 DB 3 LALLEDCWCRMSVDEQKSLMVTGIPADFEAEIQEVLQETLKSIGRYRLGKIFRKQENA 62
 QY 62 NAVLELLEDDTDSAIPTSEVQGGVWKVIFKTPNQDTFELRLNLFLEKGTQVSGMFR 121
 DB 63 NAVLELLEDDTDSAIPTSEVQGGVWKVIFKTPNQDTFELRLNLFLEKGTQVSGMFR 122
 QY 122 ALGOEALSPATVPCISPELLAHLLGQMAHAPQLPMPYRKLRFVSGSAVPAPEESFE 181
 DB 123 ALGOEGVSPATVPCISPELLAHLLGQMAHAPQLPMPYRKLRFVSGSAVPAPEESFE 182
 QY 182 VWLEQATEIVKEWP 195
 DB 183 VWLEQATEIVKEWP 196
 RESULT 2
 PMA2 MACFA STANDARD; PRT; 364 AA.
 ID _PMA2 MACFA
 AC Q9G0U3; 2004 (Rel. 43, Created)
 DT 15-MAR-2004 (Rel. 43, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Paraneoplastic antigen Ma2 homolog.
 GN PNMA2.
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecinae; Macaca.
 OX NCBI_TaxID=9541;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
 RA Suzuki Y., Sugano S., Hashimoto K.;
 RT "Isolation of full-length cDNA clones from macaque brain cDNA
 RT libraries";
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: Nuclear; nucleolar (By similarity).
 CC -1- SIMILARITY: Belongs to the PNMA family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL; AB047632; BAB12156.1; -
 KW Nuclear protein.

FT DOMAIN 333 338 POLY-GLU.
 SQ SEQUENCE 364 AA; 41350 MW; 0CF72210D7EC1524 CRC64;
 Query Match 96.3%; Score 959; DB 1; Length 364;
 Best Local Similarity 95.9%; Pred. No. 2.5e-74;
 Matches 186; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
 QY 2 LALLEDCWCRMSVDEQKSLMVTGIPADFEAEIQEVLQETLKSIGRYRLGKIFRKQENA 61
 DB 3 LALLEDCWCRMSVDEQKSLMVTGIPADFEAEIQEVLQETLKSIGRYRLGKIFRKQENA 62
 QY 62 NAVLELLEDDTDSAIPTSEVQGGVWKVIFKTPNQDTFELRLNLFLEKGTQVSGMFR 121
 DB 63 NAVLELLEDDTDSAIPTSEVQGGVWKVIFKTPNQDTFELRLNLFLEKGTQVSGMFR 122
 QY 122 ALGOEALSPATVPCISPELLAHLLGQMAHAPQLPMPYRKLRFVSGSAVPAPEESFE 181
 DB 123 ALGOEGVSPATVPCISPELLAHLLGQMAHAPQLPMPYRKLRFVSGSAVPAPEESFE 182
 QY 182 VWLEQATEIVKEWP 195
 DB 183 VWLEQATEIVKEWP 196
 RESULT 3
 PMA2 MOUSE STANDARD; PRT; 365 AA.
 ID _PMA2 MOUSE
 AC Q9BHK0; 2004 (Rel. 43, Created)
 DT 15-MAR-2004 (Rel. 43, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Paraneoplastic antigen Ma2 homolog.
 GN PNMA2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Brain cortex;
 RA MEDLINE=22354683; PubMed=12466851;
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
 RA Nikaide I., Oeato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojibori T.,
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
 RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
 RA Blake J.A., Bradt D., Brusci V., Chothia C., Corbani L.E., Cousins S.,
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
 RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
 RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
 RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
 RA Sandelin A., Schneider C., Sempile C.A., Setou M., Shimada K.,
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
 RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
 RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
 RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
 RA Shiraki T., Waki K., Kawai K., Imotani K., Ishii Y., Itoh M., Kagawa I.,
 RA Hara A., Hashizume W., Imotani K., Shibata K., Shinagawa A.,
 RA Miyazaki A., Sakai K., Sasaki D., Waterston R., Lander E.S., Rogers J.,
 RA Yasunishi A., Yoshino M., Yoshizaki Y.;
 RA Birney E., Hayashizaki Y.;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs";
 RL Nature 420:563-573(2002).
 CC -1- SUBCELLULAR LOCATION: Nuclear; nucleolar (By similarity).
 CC -1- SIMILARITY: Belongs to the PNMA family.

1. The first step in the process of identifying a problem is to recognize that a problem exists. This involves gathering information about the situation and identifying the specific issue that needs to be addressed.

NCBI_TaxID=10090;
[1]
SEQUENCE FROM N.A.
STRAIN=C57BL/6J; TISSUE=Brain, and Embryo;
MEDLINE=22354683; PubMed=12466851;
OKazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kiyosawa H.,
RA Nikola I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojohori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Gough J.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawai H., Kawasawa Y., Kedzierzki R.M., King B.L.,
RA Konegaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavalon M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Birney E., Hayashizaki Y.,
RA "Analysis of the mouse transcriptome based on functional annotation of
RT Nature 420:563-573(2002)."
RL 60,770 full-length cDNAs."
CC -!- SUBCELLULAR LOCATION: Nuclear; nucleolar (By similarity).
CC -!- SIMILARITY: Belongs to the PNMA family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; AK017476; BAB30762.1; -;
DR EMBL; AK028331; BAC25895.1; -;
DR MGI; MGI:2180564; Pnmal.
KW Nuclear protein.
FT DOMAIN 336 341 POLY-GLU.
FT CONFLICT 136 136 T -> M (IN REF. 1; BAB30762).
SQ SEQUENCE 353 AA; 39688 MW; 18CEDC3AC4E70939 CRC64;

Query Match 46.1%; Score 459; DB 1; Length 353;
Best Local Similarity 47.9%; Pred. No. 8.9e-32;
Matches 93; Conservative 36; Mismatches 59; Indels 6; Gaps 4;

QY 2 LALLEDCWCRIMSVDEQKSLMTVGIPADFEAEAEIQVLETLKSLGRYLLGKIFKQENA 61
Db 3 MTLLEDWCRGMDVNSQALLVWGIPVNCDEAEIETLQAAMPQVS-YRVLGRMFWEENA 61

QY 62 NAVLLELLEDDTVSAIPSEVQGGVWVKVIFKTPNQDTEFLRLNLFLEKEGQTVSGMFR 121
Db 62 KAALLELTGAVDYAALPREMPGKGGVWVKVIFKTPPTSDAEFLRLHLFLAREGWTVDVAR 121

QY 122 ALGQEAALSPATVPCISPELLAHLGQAMAHAPQLL-PMRYKRLRVFSSAVPAPEESF 180
Db 122 VLGFQ--NPPTP--GPENPAEMLNVLNDVQLPLVESIWKYKLTFLSGRDIPGGEETF 177

QY 181 EYVLEQATEIVKEW 194
Db 178 DPWLEHTNEVLEW 191

RESULT 6
ID PNMA1 MOUSE STANDARD; PRT; 353 AA.
AC O8C1C8; Q9CVP2;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Paraneoplastic antigen Mal homolog.
GN PNMA1 OR MAL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

Schmerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences.";
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[4]
SEQUENCE OF 162-353 FROM N.A.
TISSUE=Testis;
RC Pouskka A., Wellenreuther R., Mewes H.-W., Weil B., Wiemann S.;
RA Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Nuclear; nucleolar. In tumor cells, it is
CC cytoplasmic.
CC -!- TISSUE SPECIFICITY: Testis and brain specific. In some patients
CC suffering from cancers, it is also specifically expressed by the
CC paraneoplastic tumor cells.
CC -!- MISCELLANEOUS: Antibodies against PNMA1 are present in sera from
CC patients suffering of paraneoplastic neurological disorders.
CC -!- SIMILARITY: Belongs to the PNMA family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; AF037364; AAD13810.3; -;
DR EMBL; AF320308; AAO5100.1; -;
DR EMBL; BC039577; AAH39577.1; -;
DR EMBL; AL834327; CAD38985.1; -;
DR Genbank; HGNC:9158; PNMA1.
DR MIM; 604010; -;
DR GO; GO:0005737; C:cytoplasm; TAS.
DR GO; GO:0005730; C:nucleolus; TAS.
DR GO; GO:0007417; P:central nervous system development; TAS.
DR GO; GO:0007283; P:spermatogenesis; TAS.
DR Antigen; Tumor antigen; Nuclear protein.
KW DOMAIN 335 341 POLY-GLU.
FT CONFLICT 335 341 POLY-GLU.
SQ SEQUENCE 353 AA; 39761 MW; EB7F5B6AEDA25961 CRC64;

Query Match 46.8%; Score 466; DB 1; Length 353;
Best Local Similarity 47.9%; Pred. No. 2.3e-32;
Matches 93; Conservative 38; Mismatches 57; Indels 6; Gaps 4;

QY 2 LALLEDCWCRIMSVDEQKSLMTVGIPADFEAEAEIQVLETLKSLGRYLLGKIFKQENA 61
Db 3 MTLLEDWCRGMDVNSQALLVWGIPVNCDEAEIETLQAAMPQVS-YRVLGRMFWEENA 61

QY 62 NAVLLELLEDDTVSAIPSEVQGGVWVKVIFKTPNQDTEFLRLNLFLEKEGQTVSGMFR 121
Db 62 KAALLELTGAVDYAALPREMPGKGGVWVKVIFKTPPTSDAEFLRLHLFLAREGWTVDVAR 121

QY 122 ALGQEAALSPATVPCISPELLAHLGQAMAHAPQLL-PMRYKRLRVFSSAVPAPEESF 180
Db 122 VLGFQ--NPPTP--GPENPAEMLNVLNDVQLPLVESIWKYKLTFLSGRDIPGGEETF 177

QY 181 EYVLEQATEIVKEW 194
Db 178 DPWLEHTNEVLEW 191

RESULT 6
ID PNMA1 MOUSE STANDARD; PRT; 353 AA.
AC O8C1C8; Q9CVP2;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Paraneoplastic antigen Mal homolog.
GN PNMA1 OR MAL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

RESULT 7
MOPI HUMAN STANDARD; PRT; 351 AA.
AC Q96BY2: Q9H833; Q9HAS1.
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Modulator of apoptosis 1 (MAP-1) (Paraneoplastic antigen MA4).
GN MOAP1 OR PNMA4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A., CHARACTERIZATION, AND MUTAGENESIS OF LEU-120;
RP 120-LEU-ARG-127 AND 125-GLY-GLU-127.
RC TISSUE=Cerebellum;
RX MEDLINE=21264738; PubMed=11060313;
RA Tan K.O., Tan K.M.L., Chan S.-L., Yee K.S.-Y., Bevoit M., Ang K.C.,
RA Yu V.C.;
RT "MAP-1, a novel proapoptotic protein containing a BH3-like motif that
RT associates with Bax through its Bcl-2 homology domains.";
RL J. Biol. Chem. 276:2802-2807 (2001).
RN [2]
SEQUENCE FROM N.A.
RC TISSUE=Retinoblastoma;
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sato H., Sugano S., Shiratori A., Sudo H.,
RA Wagatsuma M., Hosoi R., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahara K., Masuho Y.,
RA Ninomiya K., Iwayanagi T.;
RT "NEDD human cDNA sequencing project.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
SEQUENCE FROM N.A.
RC TISSUE=Skin;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.P., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzyzinski M.I., Skalska U., Snailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
CC -!- FUNCTION: Mediates caspase-dependent apoptosis.
CC -!- SUBUNIT: Homodimer. Binds BAX, BCL2 and BCLX.
CC -!- TISSUE SPECIFICITY: Ubiquitously expressed with higher levels in
CC heart and brain.
CC -!- DOMAIN: The BH3-like domain is required for association with BAX
CC and for mediating apoptosis. The three BH domains (BH1, BH2, and
CC BH3) of BAX are all required for mediating protein-protein
CC interaction.
CC -!- SIMILARITY: Belongs to the PUMA family.
CC -!- CAUTION: Ref.2 sequence differs from that shown due to a
CC frameshift in position 102.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF305550; AAC31786.1; -;
DR EMBL: AK024029; BAB14788.1; ALT_SEQ.
DR EMBL: BC015044; AAH15044.1; -;
DR Genew: HGNC:16658; MOAP1.
KW Apoptosis.
FT SITE 120 127 BH3-LIKE.
FT MUTAGEN 120 127 MISSING: ABROGATED INTERACTION WITH BAX,
FT MUTAGEN 120 127 RESULTING IN A NONAPOPTOTIC PROTEIN.
FT MUTAGEN 120 120 L->E: WEAKENED INTERACTION WITH BAX,
FT MUTAGEN 125 127 RESULTING IN A NONAPOPTOTIC PROTEIN.
FT MUTAGEN 125 127 GHE->VLA: ABROGATED INTERACTION WITH BAX,
FT MUTAGEN 125 127 RESULTING IN A NONAPOPTOTIC PROTEIN.
FT CONFLICT 244 244 T -> A (IN REF. 2).
FT CONFLICT 258 258 Y -> H (IN REF. 2).
FT CONFLICT 259 259 Q -> H (IN REF. 1).
SQ SEQUENCE 351 AA; 5310142AC02B563C CRC64;
Query Match 43.8%; Score 436; DB 1; Length 351;
Best Local Similarity 46.9%; Pred. No. 8.1e-30;
Matches 91; Conservative 30; Mismatches 71; Indels 2; Gaps 2;
QY 2 LALLDWCRCMSVDEQKSLMVTGIPADFEAEATQEVLTQSLGRLGKIFRQENA 61
Db LALLDWCRCMSVDEQKSLMVTGIPADFEAEATQEVLTQSLGRLGKIFRQENA 61
QY 3 LRLLDWCRCMDMNPRAKLLIAGISOSCSVAEEALQAGLAPLGEVRLGMRFRDNR 62
Db LRLLDWCRCMDMNPRAKLLIAGISOSCSVAEEALQAGLAPLGEVRLGMRFRDNR 62
QY 62 NAVILLELDDTVSAIPSEVQGGVYKFTPNQDTFLERLNLFLKEGQTVSGMER 121
Db NAVILLELDDTVSAIPSEVQGGVYKFTPNQDTFLERLNLFLKEGQTVSGMER 121
QY 63 KVALVGLTAETSHALVPEKIPGKGIWRVIFKPPDPDNTFLSLNFLAGMTVGLSR 122
Db KVALVGLTAETSHALVPEKIPGKGIWRVIFKPPDPDNTFLSLNFLAGMTVGLSR 122
QY 122 ALGOEALSPATVPCISPELLAHLLGQMAHAPOPLP-MRYRKLRFVSSGAVPAPEESF 180
Db ALGOEALSPATVPCISPELLAHLLGQMAHAPOPLP-MRYRKLRFVSSGAVPAPEESF 180
QY 123 ALGHENGLDPEQGMIPENWAPMLAQAAL-EALQALQCLKYLKLVFSGRESPEEGEEF 181
Db ALGHENGLDPEQGMIPENWAPMLAQAAL-EALQALQCLKYLKLVFSGRESPEEGEEF 181
QY 181 EYVLEQATEIVKVEW 194
Db EYVLEQATEIVKVEW 194
QY 182 GRWHEHTTQMIKAW 195
Db GRWHEHTTQMIKAW 195
RESULT 8
MOPI MACFA STANDARD; PRT; 351 AA.
AC Q95K14;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Modulator of apoptosis 1 (MAP-1).
GN MOAP1.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
OC Cercopitheidae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
SEQUENCE FROM N.A.
RC TISSUE=Temporal cortex;
RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
RA Suzuki Y., Sugano S., Hashimoto K.;
RT "Isolation of full-length cDNA clones from macaque brain CDNA
RT libraries".
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Mediates caspase-dependent apoptosis (By similarity).
CC -!- SUBUNIT: Homodimer. Binds BAX, BCL2 and BCLX (By similarity).
CC -!- DOMAIN: The BH3-like domain is required for association with BAX
CC and for mediating apoptosis. The three BH domains (BH1, BH2, and
CC BH3) of BAX are all required for mediating protein-protein
CC interaction (By similarity).

CC 1- SIMILARITY: Belongs to the PNMA family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AB060854; BAB46873.1; -;
 CC
 CC K W Apoptosis. 120 127 BH3-LIKE.
 CC FT SITE 336 339 POLY-GLU.
 CC SQ SEQUENCE 351 AA; 39623 MW; C7530E4496A6FFB3 CRC64;
 CC
 CC Query Match 43.3%; Score 431; DB 1; Length 351;
 CC Best Local Similarity 46.4%; Pred. No. 2.le-29;
 CC Matches 90; Conservative 30; Mismatches 72; Indels 2; Gaps 2;
 CC
 CC QY 2 LALLEDCWCRIMSVDEQKSLMTGIPADFEAEIOEVLQETLSLGRVRLGKIFRQENA 61
 CC Db ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
 CC 3 LRLLEDWCRCMDNPKRALLIAGTISQSCVAEIEALQGLAPLGEYRLGLRFRDRDENR 62
 CC QY 62 NAVLELLEDDTDSVAIPSEVQGGVWKVIFKTPNQDTFELRLNLFLEKGGQTVSGMFR 121
 CC Db ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
 CC 63 KVALVGLTAETSHALVPRKIPGKGIVRWVIFKPPDSNTFLSLRNEFLAGEGWTGELTR 122
 CC QY 122 ALGOEALSPATVPCISPELLAHLGQMAHAPOQLP-MRYRKLRFVSGSAVPAPBESF 180
 CC Db ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
 CC 123 ALAHENSLDLEQGMIFEMWAPMLAQL-BALQALQCLKYKLRVFGREPPPGGEEF 181
 CC QY 181 EVMLEQATEIVKEW 194
 CC Db ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
 CC 182 GRWFFHTTOMIKAW 195
 CC
 CC RESULT 9
 CC MOPI MOUSE STANDARD; PRT; 352 AA.
 CC ID MOPI MOUSE
 CC AC Q9ERH6;
 CC DT 15-MAR-2004 (Rel. 43, Created)
 CC DT 15-MAR-2004 (Rel. 43, Last sequence update)
 CC DT 15-MAR-2004 (Rel. 43, Last annotation update)
 CC DE Modulator of apoptosis 1 (MAP-1).
 CC GN MOAP1.
 CC OS Mus musculus (Mouse).
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CC OX NCBI_TaxID=10090;
 CC [1]
 CC RN SEQUENCE FROM N.A.
 CC RC TISSUE=Brain;
 CC RX MEDLINE=21264738; PubMed=11060313;
 CC RA Tan K.O., Tan K.M.L., Chan S.-L., Yee K.S.Y., Bevoit M., Ang K.C.,
 CC RA Yu V.C.;
 CC RT "MAP-1, a novel proapoptotic protein containing a BH3-like motif that
 CC RT associates with Bax through its Bcl-2 homology domains."
 CC RL J. Biol. Chem. 276:2802-2807(2001).
 CC [2]
 CC RN SEQUENCE FROM N.A.
 CC RC STRAIN=C57BL/6J; TISSUE=Testis;
 CC RX MEDLINE=21085660; PubMed=11217851;
 CC RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 CC RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 CC RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.I.,
 CC RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 CC RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 CC RA Fleischmann W., Gaasterland T., Giasi C., King B., Kochiwa H.,
 CC RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
 CC RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 CC RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 CC RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,

RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
 RA Hayashizaki Y.;
 RA "Functional annotation of a full-length mouse cDNA collection."
 RT Nature 409:685-690(2001).
 RL [3]
 RN SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6NCR; TISSUE=Hematopoietic, and Kidney;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner K.H., Schaefer C.F., Bhat N.K.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins S.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong E.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Vallaloo D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whitig M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RA "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC 1- FUNCTION: Mediates caspase-dependent apoptosis.
 CC 1- SUBUNIT: Homodimer. Binds BAX, BCL2 and BCLX (By similarity).
 CC 1- DOMAIN: The BH3-like domain is required for association with BAX
 CC and for mediating apoptosis. The three BH domains (BH1, BH2, and
 CC BH3) of BAX are all required for mediating protein-protein
 CC interaction (By similarity).
 CC 1- SIMILARITY: Belongs to the PNMA family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF305551; AAG31787.1; -;
 CC DR EMBL; AK019599; BAB31810.1; -;
 CC DR EMBL; BC014715; AAI14715.1; -;
 CC DR EMBL; BC055374; AAI55374.1; -;
 CC DR MGI; MGI:1915555; Mgap1.
 CC KW Apoptosis.
 CC FT SITE 120 127 BH3-LIKE.
 CC FT DOMAIN 335 340 POLY-GLU.
 CC FT CONFLICT 57 57 R -> K (IN REF. 3; AAF55374).
 CC SQ SEQUENCE 352 AA; 39404 MW; 8F4630D080495D98 CRC64;
 CC
 CC Query Match 43.3%; Score 431; DB 1; Length 352;
 CC Best Local Similarity 44.9%; Pred. No. 2.2e-29;
 CC Matches 89; Conservative 35; Mismatches 66; Indels 8; Gaps 3;
 CC
 CC QY 2 LALLEDCWCRIMSVDEQKSLMTGIPADFEAEIOEVLQETLSLGRVRLGKIFRQENA 61
 CC Db ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
 CC 3 LRLLEDWCRCMDNPKRALLIAGTISQSCVAEIEALQGLAPLGEYRLGLRFRDRDENR 62
 CC QY 62 NAVLELLEDDTDSVAIPSEVQGGVWKVIFKTPNQDTFELRLNLFLEKGGQTVSGMFR 121
 CC Db ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
 CC 63 NVALIGLTVETGSLVPRKIPAGKGVWRVIFKPPDSDFLQRLNLFLEKGGQTVSGMFR 122

QY 122 ALGQE-----ALSPATVFCISPELLAHLGQAMAHAPOLLP-MRYRKLRFVSGSAVPE 176
 Db 123 VLGNRPDLGLDPG---IMPEIRAPMLAQALNEALPTLQYLRKLSVSGRDPGPG 179
 QY 177 EESPEVWLEQATEIVKSW 194
 Db 180 EEFESWMTHTSQVMKTW 197

RESULT 10
 CU05_HUMAN
 ID CU05_HUMAN STANDARD; PRT; 2298 AA.
 AC Q9Y3R5; Q9Y3R5;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Protein C21orf5.
 GN C21ORF5 OR KIAA0933.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20408884; PubMed=10950924;
 RA Guipponi M., Brunschwig K., Chamoun Z., Scott H.S., Shibuya K.,
 RA Kudoh J., Delezoide A.L., El Samadi S., Chettouh Z., Rossier C.,
 RA Shimizu N., Mueller F., Delabar J.M., Antonarakis S.E.;
 RT "C21orf5, a novel human chromosome 21 gene, has a Caenorhabditis
 RT elegans ortholog (pad-1) required for embryonic patterning.";
 RL Genomics 68:30-40 (2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Shibuya K., Kudoh J., Minoshima S., Kawasaki K., Nakatoh E.,
 RA Shintani A., Asakawa S., Shimizu N.;
 RT "Genomic sequencing of 1.2-Mb region on human chromosome 21q22.2.";
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20289799; PubMed=10830953;
 RA Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,
 RA Park H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Groner Y.,
 RA Soeda E., Onki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K.,
 RA Polley A., Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,
 RA Reichwald A., Rump A., Schillhabel M., Schudy A., Zimmermann W.,
 RA Rosenthal K., Kudoh J., Shibuya K., Kawasaki K., Asakawa S.,
 RA Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E.,
 RA Minoshima S., Shimizu N., Nordieck G., Hornischer K., Brandt P.,
 RA Scharfe M., Schoen O., Desario A., Reichelt J., Kauer G., Bloecker H.,
 RA Ramser J., Beck A., Klages S., Hennig S., Riessmann L., Dagand E.,
 RA Wehrmeyer S., Borzym K., Gardiner K., Nizetic D., Francis F.,
 RA Lehrach H., Reinhardt R., Yaspo M.-L.;
 RT "The DNA sequence of human chromosome 21.";
 RL Nature 405:311-319 (2000).
 RN [4]
 RP SEQUENCE OF 152-2298 FROM N.A.
 RX TISSUE=Brain;
 RC MEDLINE=22158633; PubMed=12168954;
 RA Nakajima D., Okazaki N., Yamakawa H., Kikuno R., Ohara O., Nagase T.;
 RT "Construction of expression-ready cDNA clones for KIAA genes: manual
 RT curation of 330 KIAA cDNA clones.";
 RL DNA Res. 9:99-106 (2002).
 RN [5]
 RP SEQUENCE OF 677-2298 FROM N.A.
 RX TISSUE=Brain;
 RC MEDLINE=99246063; PubMed=10231032;
 RA Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirokawa M.,
 RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. XIII.
 RT The complete sequences of 100 new cDNA clones from brain which code
 RT for large proteins in vitro.";
 RL DNA Res. 6:63-70 (1999).
 CC -!- TISSUE SPECIFICITY: Ubiquitous.

CC -!- SIMILARITY: TO C.ELEGANS Y18D10A.15.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; AJ237839; CAB41415.1; -.
 DR EMBL; AP000692; BAA89431.1; -.
 DR EMBL; AP000699; BAA89431.1; JOINED.
 DR EMBL; AP000890; BAA89431.1; JOINED.
 DR EMBL; AP000891; BAA89431.1; JOINED.
 DR EMBL; AP001725; BAA95548.1; -.
 DR EMBL; AB023150; BAA76777.2; -.
 DR Genew; HGNC:1291; C21orf5.
 DR MIM; 604803; -.
 DR GO; GO:0007275; P:development; NAS.
 DR InterPro; IPR007249; Dopey_N.
 DR Pfam; PF04118; Dopey_N; 1.
 DR FT CONFLICT 498 498 Q -> H (IN REF. 1).
 FT CONFLICT 1118 1118 G -> C (IN REF. 2 AND 3).
 FT CONFLICT 1149 1149 H -> P (IN REF. 2 AND 3).
 FT CONFLICT 1318 1318 L -> P (IN REF. 1).
 FT CONFLICT 1886 1886 V -> A (IN REF. 1).
 FT CONFLICT 1974 1974 E -> G (IN REF. 1).
 FT CONFLICT 2139 2139 G -> E (IN REF. 4).
 FT CONFLICT 2154 2154 K -> R (IN REF. 1).
 SQ SEQUENCE 2298 AA; 258221 MW; 2C49AF5550739F0 CRC64;

Query Match 9.1%; Score 91; DB 1; Length 2298;
 Best Local Similarity 22.7%; Pred. No. 19;
 Matches 53; Conservative 33; Mismatches 89; Indels 58; Gaps 11;

QY 1 PLALLE-----DWCEIMSVDEQKSLMWVGIPADPEEAIEQVLTSLKSLGRYLLGKI 54
 Db 1579 PUTLLEGLTTISHFCLLEQANQKTKMAAGDPANLRNAR-NAILELPRTVTMTALLWNV 1637
 QY 55 FRKQENANAVILLELELTDVSAIPSEVOGKGVKVIKTPNQD-TEFLERLNLFL---- 109
 Db 1638 LRKEHTQKRP-VDLLGAT-----KGSSVYFKTKTIQKILDFLNPJTAHLGVOL 1687
 QY 110 -----EKEGQTVSGMFRALGQALSPAT-----VPCISPELLAHLGQAMAHAP 153
 Db 1688 TAAVAAVMSRKAQRHSM-KIIPITASASQLFLDLVLCALSTLQTDTLHLKVEVVKRP 1746
 QY 154 Q-----PLLPWRYKRLRVFSGSAVPAPE-BESFEVWLEQATEIVK 192
 Db 1747 QVKGDEKSLVDIPVLOFQYAFLO-----RLFPALQENFSSLLGVLKESVQ 1794

RESULT 11
 TRAP_HUMAN
 ID TRAP_HUMAN STANDARD; PRT; 3859 AA.
 AC Q9Y4A5; Q9Y4A5; Q9Y6H4;
 DT 15-MAR-2004 (Rel. 43, Created)
 DT 15-MAR-2004 (Rel. 43, Last sequence update)
 DE Transcription/transcription domain-associated protein (350/400 kDa
 DE PCAP-associated factor) (PAF350/400) (STAF40) (Tral homolog).
 GN TRRAP OR PAF400.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 2), PARTIAL SEQUENCE, FUNCTION,
 RP SUBCELLULAR LOCATION, AND INTERACTION WITH MYC AND E2F1.
 RC TISSUE=Cervical carcinoma;
 RX MEDLINE=98372443; PubMed=9708738;
 RA McMahon S.B., Van Buskirk H.A., Dugan K.A., Copeland T.D., Cole M.D.;

RT "The novel ATM-related protein TRRAP is an essential cofactor for the
RT c-Myc and E2F oncoproteins.";
RL Cell 94:363-374(1998).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1), SEQUENCE OF 41-90; 337-344; 369-387;
RP 879-892; 997-1005; 1171-1184; 1237-1247; 1545-1560; 1815-1820;
RP 1922-1934; 2211-2218; 2260-2275; 2534-2547; 2583-2594; 2706-2726;
RP 2830-2844; 3567-3573; 3583-3598; 3604-3614; 3712-3730 AND 3822-3834,
RP AND IDENTIFICATION IN THE PCAF COMPLEX WITH TADA2L; TADA3L; TAF5L;
RP TAF6L; TAF10; SUPT3H; TAF12 AND TAF9.
RC TISSUE=Petal heart;
RX MEDLINE=99102960; PubMed=9885574;
RA Vassiliev A., Yamauchi J., Kotani T., Prives C., Avantaggiati M.L.,
RA Qin J., Nakatani Y.;
RA "The 400 kDa subunit of the PCAF histone acetylase complex belongs to
RT the ATM superfamily.";
RL Mol. Cell 2:869-875(1998).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=22737999; PubMed=12853948;
RA Hillier L.W., Fulton R.S., Fulton L.A., Graves T.A., Pepin K.H.,
RA Wagner-McPherson C., Layman D., Maas J., Jaeger S., Walker R.,
RA Wylie K., Sekhon M., Becker M.C., O'Laughlin M.D., Schaller M.E.,
RA Fewell G.A., Delehaanty K.D., Miner T.L., Nash W.E., Cordes M., Du H.,
RA Sun H., Edwards J., Bradshaw-Cordum H., Ali J., Andrews S., Isak A.,
RA Vanbrunt A., Nguyen C., Du F., Lamar B., Courtney L., Kalicki J.,
RA Ozerky P., Bielicki L., Scott K., Holmes A., Harkins R., Harris A.,
RA Strong C.M., Hou S., Tomlinson C., Dauphin-Kohlberg S.,
RA Tin-Lowicz A.-M., Abbott A., Minx P., Maupin R., Stromatt C.,
RA Latrelle P., Miller N., Johnson D., Murray J., Woessner J.P.,
RA Wendt M.C., Yang S.-P., Schultz B.R., Wallis J.W., Spieth J.,
RA Bieri T.A., Nelson J.O., Berkowicz N., Wohldmann P.E., Cook L.L.,
RA Hickenbotham M.T., Eldred J., Williams D., Bedell J.A., Mardis E.R.,
RA Clifton S.W., Chissoe S.L., Marra M.A., Raymond C., Haugen E.,
RA Gillett W., Zhou Y., James R., Phelps K., Iadamoto S., Bubbs K.,
RA Simms E., Levy R., Clendenning J., Kaul R., Kent W.J., Purey T.S.,
RA Baerbach R.A., Brent M.R., Keibler E., Flicek P., Bork P., Suyama M.,
RA Bailey J.A., Portnoy M.E., Torrents D., Chinwalla A.T., Gish W.R.,
RA Eddy S.R., McPherson J.D., Olson M.V., Eichler E.E., Green E.D.,
RA Waterston R.H., Wilson R.K.;
RA "The DNA sequence of human chromosome 7.";
RL Nature 424:157-164(2003).
RN [4]
RP SEQUENCE OF 21-38; 882-893; 2005-2015; 2049-2059; 2063-2074 AND
RP 3583-3598, AND IDENTIFICATION IN THE STAGA COMPLEX WITH SF3B3;
RP GCN5L2; KIAA0764; TAF5L; TAF6L; TADA3L; TAF10; TAF12 AND TAF9.
RX MEDLINE=21448975; PubMed=11564863;
RA Martinez E., Palhan V.B., Tjernerberg A., Lymar E.S., Gamper A.M.,
RA Kundu T.K., Chait B.T., Roeder R.G.;
RA "Human STAGA complex is a chromatin-acetylating transcription
RT coactivator that interacts with pre-mRNA splicing and DNA
RT damage-binding factors in vivo.";
RL Mol. Cell. Biol. 21:6782-6795(2001).
RN [5]
RP IDENTIFICATION IN THE TFC-HAT COMPLEX WITH TAF5L; TAF6L; TADA3L;
RP SUPT3H; TAF2; TAF4; TAF5; GCN5L2 AND TAF10.
RX MEDLINE=99303588; PubMed=10373431;
RA Brand M., Yamamoto K., Staub A., Tori L.;
RA "Identification of TATA-binding protein-free TAFII-containing complex
RT subunits suggests a role in nucleosome acetylation and signal
RT transduction.";
RL J. Biol. Chem. 274:18285-18289(1999).
RN [6]
RP IDENTIFICATION IN THE TIP60 HAT COMPLEX WITH HTATIP; RUVBL1 AND
RP RUVBL2.
RX MEDLINE=20419292; PubMed=10966108;
RA Ikura T., Ogryzko V.V., Grigoriev M., Groisman R., Wang J.,
RA Horikoshi M., Scully R., Qin J., Nakatani Y.;
RA "Involvement of the TIP60 histone acetylase complex in DNA repair and
RT apoptosis.";
RL Cell 102:463-473(2000).
RN [7]

RP INTERACTION WITH GCN5L2.
RX MEDLINE=20079268; PubMed=10611234;
RA McMahon S.B., Wood M.A., Cole M.D.;
RT "The essential cofactor TRRAP recruits the histone acetyltransferase
RT hGCN5 to c-Myc.";
RL Mol. Cell. Biol. 20:556-562(2000).
RN [8]
RP INTERACTION WITH E2F1 AND E2F4, AND FUNCTION.
RX MEDLINE=21413878; PubMed=11418595;
RA Lang S.E., McMahon S.B., Cole M.D., Hearing P.;
RT "E2F transcriptional activation requires TRRAP and GCN5 cofactors.";
RL J. Biol. Chem. 276:32627-32634(2001).
RN [9]
RP DOMAIN.
RX MEDLINE=21338486; PubMed=11445536;
RA Park J., Kunjibettu S., McMahon S.B., Cole M.D.;
RT "The ATM-related domain of TRRAP is required for histone
RT acetyltransferase recruitment and Myc-dependent oncogenesis.";
RL Genes Dev. 15:1619-1624(2001).
RN [10]
RP FUNCTION, AND INTERACTION WITH TP53.
RX MEDLINE=22133580; PubMed=12138177;
RA Ard P.G., Chatterjee C., Kunjibettu S., Adside L.R., Gralinski L.E.,
RA McMahon S.B.;
RT "Transcriptional regulation of the mdm2 oncogene by p53 requires TRRAP
RT acetyltransferase complexes.";
RL Mol. Cell. Biol. 22:5650-5661(2002).
RN [11]
RP IDENTIFICATION IN THE BAF53 COMPLEX WITH BAF53A; RUVBL1 AND SMARCA4.
RX MEDLINE=21829261; PubMed=11839798;
RA Park J., Wood M.A., Cole M.D.;
RT "BAF53 forms distinct nuclear complexes and functions as a critical
RT c-Myc-interacting nuclear cofactor for oncogenic transformation.";
RL Mol. Cell. Biol. 22:1307-1316(2002).
RN [12]
RP FUNCTION.
RX MEDLINE=22628784; PubMed=12743606;
RA Lang S.E., Hearing P.;
RT "The adenovirus E1A oncoprotein recruits the cellular TRRAP/GCN5
RT histone acetyltransferase complex.";
RL Oncogene 22:2836-2841(2003).
RN [13]
RP FUNCTION.
RX MEDLINE=22651123; PubMed=12660246;
RA Liu X., Tesfai J., Evarad Y.A., Dent S.Y.R., Martinez E.;
RT "c-Myc transformation domain recruits the human STAGA complex and
RT requires TRRAP and GCN5 acetylase activity for transcription
RT activation.";
RL J. Biol. Chem. 278:20405-20412(2003).
RN [14]
RP FUNCTION: Adapter protein, which is found in various multiprotein
CC chromatin complexes with histone acetyltransferase activity (HAT),
CC which gives a specific tag for epigenetic transcription
CC activation. Plays a central role in MYC (c-Myc) transcription
CC activation, and also participates in cell transformation by MYC.
CC Required for TP53/p53-, E2F1- and E2F4-mediated transcription
CC activation. Also involved in transcription activation mediated by
CC the adenovirus E1A, a viral oncoprotein that deregulates
CC transcription of key genes. Probably acts by linking transcription
CC factors such as E1A, MYC or E2F1 to HAT complexes such as STAGA
CC thereby allowing transcription activation. Probably not required
CC in the steps following histone acetylation in processes of
CC transcription activation. May be required for the mitotic
CC checkpoint and normal cell cycle progression.
CC
CC -!- SUBUNIT: Interacts with MYC, E2F1 and E2F4 transcription factors.
CC Interacts directly with TP53. Interacts with GCN5L2. Component of
CC various HAT complexes. Component of the PCAF complex, at least
CC composed of TADA2L/ADA2, SUPT3H, TADA3L/ADA3, TAF5L/PAF65beta,
CC TAF6L/PAF65alpha, TAF10/TAF1130, TAF12/TAF1120, TAF9/TAF1131 and
CC TRRAP. Component of the TFC-HAT complex, at least composed of
CC TAF5L, TAF6L, TADA3L, SUPT3H/SPT3, TAF2/TAF1150, TAF4/TAF1135,
CC TAF5/TAF1100, GCN5L2/GCN5, TAF10 and TRRAP. Component of the
CC TIP60 HAT complex, at least composed of HTATIP/TIP60,
CC RUVBL1/TIP49, RUVBL2/TIP48 and TRRAP, which preferentially

CC acetylates histone H4 (and H2A) within nucleosomes. Component of
 CC the STAGA complex, at least composed of SP3B3/SAP130, GCN5L2,
 CC STAF65gamma/KIAA0764, TAF5L, TAF6L, TADA3L, TAF10, TAF12, TRRAP
 CC and TAF9. Component of the BAF53 complex, at least composed of
 CC BAF53A, RUVBL1, SMARCA4/BRG1, and TRRAP, which preferentially
 CC acetylates histone H4 (and H2A) within nucleosomes.
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoId=Q9Y4A5-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=Q9Y4A5-2; Sequence=VSP_009102, VSP_009103;
 CC DOMAIN: The PI3K/PI4K domain is required for the recruitment of
 CC HAT complexes, and the MYC-dependent transactivation. Although it
 CC is strongly related to the PI3/PI4-kinase family, it lacks the
 CC typical motifs that constitute the catalytic site of PI3/PI4-
 CC kinase proteins, and lacks such activity.
 CC -!- SIMILARITY: Belongs to the PI3/PI4-kinase family. TRAL subfamily.
 CC -!- SIMILARITY: Contains 2 TPR repeats.
 CC -!- CAUTION: Ref.3 sequences differ from that shown due to erroneous
 CC gene model prediction.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL; AF076974; AAD09420.1; -;
 CC DR EMBL; AF110377; AAD04629.1; -;
 CC DR EMBL; AC004893; AAC62433.1; ALT_SEQ.
 CC DR EMBL; AC004991; AAC27675.1; ALT_SEQ.
 CC DR PIR; T02632; T02632.

Query Match 8.8%; Score 87.5; DB 1; Length 3859;
 Best Local Similarity 24.0%; Pred. No. 73;
 Matches 50; Conservative 28; Mismatches 71; Indels 59; Gaps 9;
 QY 20 LMVTGIPADFEAEIQEVLOETLKSIGR-----YRLGKIFRQENANAVLL 66
 Db 851 LCVNLDLPDQDHIQPVRAELMQALWRLNPADSIHVAYRVLGKF-----GGSNRKWL 906
 QY 67 ELLEDTDSVAIPSEVQGGVKNVIFKTPNQDTFLE---RLNLEKEGQT----- 115
 Db 907 K--ESQKLHYVTVVQGG-----PSITVFSDCKASIQLPMEKAIEITLDCLSA 953
 QY 116 -VSGMFRALGQALSPATVPCISPELLAHLGQAMAH-----APQPLLMRYRKLRFV 167
 Db 954 NTEPYRRQAEVIRKCFVLVAMNSLEDNKHAYQLLAHPNFTKTPINVIISHRYK----- 1008
 QY 168 SGSAPVAPAEESFEVWLQQA--TEIVKE 193
 Db 1009 ---AQTARKTFEQALTGAFFMSAVIKD 1033

RESULT 12
 CET1 CANAL
 ID CET1 CANAL STANDARD; PRT; 520 AA.
 AC O93803;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE mRNA capping enzyme beta subunit (Polynucleotide 5'-triphosphatase)
 DE (EC 3.1.3.33) (mRNA 5'-triphosphatase) (TPase).
 GN CET1.
 OS Candida albicans (Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
 OX NCBI_TaxID=5476;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=IFO 1060;
 RX MEDLINE=98427288; PubMed=9755857;
 RA Yamada-Okabe T., Mio T., Matsui M., Kashima Y., Arisawa M.,
 RA Yamada-Okabe H.;
 RT "Isolation and characterization of the Candida albicans gene for mRNA
 RT 5'-triphosphatase: association of mRNA 5'-triphosphatase and mRNA 5'-
 RT guanylyltransferase activities is essential for the function of mRNA
 RT 5'-capping enzyme in vivo.";
 RL FEBS Lett. 435:49-54 (1998).
 CC -!- FUNCTION: First step of mRNA capping. Converts the 5'-triphosphate
 CC end of a nascent mRNA chain into a diphosphate end.
 CC -!- CATALYTIC ACTIVITY: 5'-phosphopolynucleotide + H(2)O =
 CC polynucleotide + phosphate.
 CC -!- COFACTOR: Requires divalent ions (By similarity).
 CC -!- SUBUNIT: The mRNA capping enzyme is composed of two separate
 CC chains alpha and beta, respectively a mRNA guanylyltransferase and
 CC an RNA 5'-triphosphatase.
 CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
 CC -!- SIMILARITY: BELONGS TO THE FUNGAL TPASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL; AB016242; BAA33965.1; -;
 CC DR InterPro; IPR004206; mRNA triPase.
 CC DR Pfam; PF02940; mRNA_triPase; 1.
 CC KW Hydrolase; mRNA processing; mRNA capping; Nuclear protein.
 CC FT DOMAIN 92 95 POLY-SER.
 CC SQ SEQUENCE 520 AA; 58791 MW; 5C0690F547E13EEC CRC64;

 Query Match 8.5%; Score 84.5; DB 1; Length 520;
 Best Local Similarity 23.1%; Pred. No. 11;
 Matches 51; Conservative 39; Mismatches 78; Indels 53; Gaps 11;
 QY 7 DMCRTMSVDEQKSLMVTG-IP-----ADFEAEIQEVLOETLKSIGRLLGKIFRK 57
 Db 241 DYTTRSVDLGCS--ITGMIPSSITRKIAEWYANFSNVEEKKRNVELELKFGLIIDK 298
 QY 58 QENANAVLELLED---TDVSAIPSEVQGGVKNVIFKTPNQDTFLERLNLFLE--KE 112
 Db 299 R-SGNRIDLVNVTCEITDHSVFEDMQVEEVAWKEI-----TKFLDELEKSFQEGKK 350
 QY 113 GQTVSGM-----FRALGQEA-----LSPATVPCISPELLAHLGQAMAH 152
 Db 351 GRKFKTSLSDNTDSFYQGRKGEGHPKRIKRVTKDNLSPRLVAIQKERVADL-----YIHN 406
 QY 153 POPLLMRYRKLRFVSGSAVPAPAEESFEVWLQQAITEIVKE 193
 Db 407 PGSLSF-----DLRLSMSLEIPVPQGNIESIITKNKPEWURE 442

 RESULT 13
 SRF2_BACSU
 ID SRF2_BACSU STANDARD; PRT; 3587 AA.
 AC Q04747;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Surfactin synthetase subunit 2.
 GN SRFAB OR SRFAB2 OR COML OR BSU03490.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE OF 1-3077 FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=93181186; PubMed=8441623;

RA Puma S., Fujishima Y., Corbell N., D'Souza C., Nakano M.M.,
RA Zuber P., Yamane K.;
RT "Nucleotide sequence of 5' portion of srfa that contains the region
RT required for competence establishment in *Bacillus subtilis*.";
RL Nucleic Acids Res. 21:93-97(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=168 / JH642;
RC MEDLINE=93360813; PubMed=8355609;
RA Cosmina P., Rodriguez F., de Ferri F., Grandi G., Perego M.,
RA Venena G., van Sinderen D.;
RT "Sequence and analysis of the genetic locus responsible for surfactin
RT synthesis in *Bacillus subtilis*.";
RL Mol. Microbiol. 8:821-831(1993).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=168;
RC MEDLINE=97124189; PubMed=8969502;
RA Yamane K., Kumano M., Kurita K.;
RT "The 25 degrees-36 degrees region of the *Bacillus subtilis*
RT chromosome: determination of the sequence of a 146 Kb segment and
RT identification of 113 genes.";
RL Microbiology 142:3047-3056(1996).
RN [4]
RP SEQUENCE FROM N.A.
RX STRAIN=168;
RC MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer J., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borries R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Etian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Gim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,
RA Guisepi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
RA Halbert H., Holbappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Roche E., Roche B., Rose M., Sadale Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowaka A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpsira P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium *Bacillus*
RT *subtilis*.";
RL Nature 390:249-256(1997).
RN [5]
RP SEQUENCE OF 514-800 FROM N.A.
RX STRAIN=ATCC 21332;
RC MEDLINE=92290255; PubMed=1601288;
RA Borchert S., Patil S.S., Marahiel M.A.;
RT "Identification of putative multifunctional peptide synthetase genes
RT using highly conserved oligonucleotide sequences derived from known
RT synthetases.";
RL FEMS Microbiol. Lett. 71:175-180(1992).
CC -!- FUNCTION: THIS PROTEIN IS A MULTIFUNCTIONAL ENZYME ABLE TO
CC ACTIVATE AND POLYMERIZE THE AMINO ACIDS LEU, GLU, ASP AND VAL.
CC ACTIVATION SITES FOR THESE AA CONSIST OF INDIVIDUAL DOMAINS.
CC -!- COFACTOR: Contains 3 covalently bound phosphopantetheines.
CC -!- PATHWAY: Cyclic peptide antibiotic surfactin biosynthesis.
CC -!- SIMILARITY: Belongs to the ATP-dependent AMP-binding enzyme
CC family.
CC -!- SIMILARITY: Contains 3 acyl carrier domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC EMBL; D13262; BAA02523.1; -;
DR EMBL; X70356; CAA49817.1; -;
DR EMBL; D50453; BAA08983.1; -;
DR EMBL; Z99105; CAB12143.1; -;
DR EMBL; X65835; CAA46678.1; -;
DR PIR; I40486; I40486.
DR HSSP; P14687; IAMU.
DR Subtilisin; BG10169; srfAB.
DR InterPro; IPR000873; AMP-bind.
DR InterPro; IPR001242; Condensatn.
DR InterPro; IPR006163; Pp-bind.
DR InterPro; IPR006162; Ppantne.S.
DR Pfam; PF00501; AMP-binding; 3.
DR Pfam; PF00668; Condensation; 4.
DR Pfam; PF00550; pp-binding; 3.
DR PRINTS; PR00154; AMPBINDING.
DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; 3.
DR PROSITE; PS00455; AMP BINDING; 3.
DR PROSITE; PS00075; ACP_DOMAIN; 3.
DR Ligase; Anticibiotic biosynthesis; Phosphopantetheine; Sporulation;
KW Multifunctional enzyme; Repeat; Complete proteome.
FT REPEAT ? 2096 DOMAIN 1 (VAL-ACTIVATING).
FT REPEAT ? 2096 DOMAIN 2 (ASP-ACTIVATING).
FT REPEAT ? 2096 DOMAIN 3 (D-LEU-ACTIVATING).
FT REPEAT ? 2096 ACYL CARRIER (ACP) 1.
FT REPEAT ? 2096 ACYL CARRIER (ACP) 2.
FT REPEAT ? 2096 ACYL CARRIER (ACP) 3.
FT REPEAT ? 2096 PHOSPHOPANTHETHEINE (POTENTIAL).
FT REPEAT ? 2096 PHOSPHOPANTHETHEINE (POTENTIAL).
FT REPEAT ? 2096 PHOSPHOPANTHETHEINE (POTENTIAL).
FT REPEAT ? 2096 S -> F (IN REF. 1).
FT REPEAT ? 2096 A -> G (IN REF. 1).
FT REPEAT ? 2096 Q -> D (IN REF. 1).
FT REPEAT ? 2096 RQA -> AOG (IN REF. 1).
FT REPEAT ? 2096 A -> V (IN REF. 1).
FT REPEAT ? 2096 L -> W (IN REF. 1).
FT REPEAT ? 2096 R -> A (IN REF. 1).
FT REPEAT ? 2096 TPA -> SRP (IN REF. 1).
FT REPEAT ? 2096 MISSING (IN REF. 5).
FT REPEAT ? 2096 A -> R (IN REF. 1).
FT REPEAT ? 2096 RHV -> ETL (IN REF. 1).
FT REPEAT ? 2096 EQSIT -> DKRIS (IN REF. 5).
FT REPEAT ? 2096 M -> L (IN REF. 5).
FT REPEAT ? 2096 PL -> LV (IN REF. 1).
FT REPEAT ? 2096 N -> I (IN REF. 1).
FT REPEAT ? 2096 H -> Q (IN REF. 1).
FT REPEAT ? 2096 V -> C (IN REF. 1).
FT REPEAT ? 2096 G -> V (IN REF. 1).
FT REPEAT ? 2096 P -> R (IN REF. 1).
FT REPEAT ? 2096 G -> E (IN REF. 1).
FT REPEAT ? 2096 KRRADG -> E (IN REF. 1).
FT REPEAT ? 2096 C -> S (IN REF. 1).
FT REPEAT ? 2096 F -> K (IN REF. 1).
FT REPEAT ? 2096 T -> S (IN REF. 1).
FT REPEAT ? 2096 GAIGRVOLVDPDAFAKPTTG -> APSPGGLICWSRCIC
FT REPEAT ? 2096 ETPDNR (IN REF. 1).
FT REPEAT ? 2096 LG -> PK (IN REF. 1).
FT REPEAT ? 2096 R -> C (IN REF. 1).
FT REPEAT ? 2096 A -> V (IN REF. 1).
FT REPEAT ? 2096 ARLTP -> LRDSL (IN REF. 1).
FT REPEAT ? 2096 E -> Q (IN REF. 1).
FT REPEAT ? 2096 ATDLF -> RQICS (IN REF. 1).
FT REPEAT ? 2096 1916
FT REPEAT ? 2096 2075
FT REPEAT ? 2096 2079
FT REPEAT ? 2096 2141
FT REPEAT ? 2096 2445
FT REPEAT ? 2096 2489
FT REPEAT ? 2096 2489


```

DR PIR; I64034; I64034.
KW TIGR; H11522; -.
DR Hypothetical protein; Complete proteome.
SQ SEQUENCE 623 AA; 67677 MW; 5EE5AE8E2818C060 CRC64;

Query Match      8.4%; Score 83.5; DB 1; Length 623;
Best local Similarity 21.7%; Pred. No. 17;
Matches 33; Conservative 19; Mismatches 51; Indels 49; Gaps 6;

OY 7 DWCIRMSDEQKSLMTVGIPAD-----PFEAEIQVLQETLKSIGRYRLLG- 52
   || : : : : || : : : : || : : : : || : : : : || : : : : ||
Db 41 DWFNNVQAEFLPSILLDDAGIQPKRLNQISIAIRKLSGKVEDFSQ-LKQADGYKIIGR 99
   || : : : : || : : : : || : : : : || : : : : || : : : : ||
OY 53 -----KIPRKQENANAVLL-----ELLEDTDVSAIPSE-----VQCKGGV 87
   || : : : : || : : : : || : : : : || : : : : || : : : : ||
Db 100 CKSVAELRTIRPTENGQRILVDAYEGSTAGGFEFVADLDLITPDGGTCFVFPNNGGR 159
   || : : : : || : : : : || : : : : || : : : : || : : : : ||
OY 88 KWVIFKTPNQDTFF-----LERINLPLE 110
   || : : : : || : : : : || : : : : || : : : : || : : : : ||
Db 160 WKRLFSSLDQTDGFGVIGGVADDTTNLNAFLD 191
   || : : : : || : : : : || : : : : || : : : : || : : : : ||

RESULT 15
PHYB_TOBAC
ID -PHYB_TOBAC STANDARD; PRT; 1132 AA.
AC P29130;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Phytochrome B.
GN PHYB.
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamids; Solanales; Solanaceae; Nicotiana.
OC NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94105358; PubMed=8278560;
RA Kern R., Gasch A., Deak M., Kay S.A., Chua N.H.:
RT "phyB of tobacco, a new member of the phytochrome family.";
RL Plant Physiol. 102:1363-1364 (1993).
RN [2]
RP SEQUENCE OF 457-1132 FROM N.A.
RX MEDLINE=92361250; PubMed=1498594;
RA Lopez-Juez E., Nagatani A., Tomizawa K.I., Deak M., Kern R.,
RA Kendrick R.E., Furuya M.:
RT "The cucumber long hypocotyl mutant lacks a light-stable PHYB-like
RT phytochrome.";
RL Plant Cell 4:241-251 (1992).
CC -!- FUNCTION: Regulatory photoreceptor which exists in two forms that
CC are reversibly interconvertible by light: the Pr form that absorbs
CC maximally in the red region of the spectrum and the Pfr form that
CC absorbs maximally in the far-red region. Photoconversion of Pr in
CC Pfr induces an array of morphogenic responses, whereas
CC reconversion of Pfr to Pr cancels the induction of those
CC responses. Pfr controls the expression of a number of nuclear
CC genes including those encoding the small subunit of ribulose
CC biphosphate carboxylase, chlorophyll A/B binding protein,
CC protochlorophyllide reductase, rRNA, etc. It also controls the
CC expression of its own gene(s) in a negative feedback fashion.
CC -!- SUBUNIT: Homodimer.
CC -!- PTM: Contains one covalently linked tetrapyrrole chromophore.
CC -!- SIMILARITY: Belongs to the phytochrome family.
CC -!- SIMILARITY: Contains 2 PAS (PER-ARNT-SIM) dimerization domains.
CC -!- SIMILARITY: Contains 1 PAS-associated C-terminal (PAC) domain.
CC -!- SIMILARITY: Contains 1 histidine kinase domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; LI0114; AAA34092.1; -
DR EMBL; M65023; AAA34093.1; -
DR PIR; T03668; T03668; ATPbind_ATPase.
DR InterPro; IPR003594; ATPbind_ATPase.
DR InterPro; IPR003018; GAF.
DR InterPro; IPR003661; His_kinA_N.
DR InterPro; IPR005467; His_kinase.
DR InterPro; IPR001610; PAC.
DR InterPro; IPR000014; PAS_domain.
DR InterPro; IPR001294; Phytochrome.
DR Pfam; PF01590; GAF; 1.
DR Pfam; PF02518; HATPase_c; 1.
DR Pfam; PF00512; Hiska; 1.
DR Pfam; PF00989; PAS; 2.
DR Pfam; PF00360; Phytochrome; 1.
DR PRINTS; PR01033; PHYTOCHROME.
DR SMART; SM00065; GAF; 1.
DR SMART; SM00387; HATPase_c; 1.
DR SMART; SM00388; Hiska; 1.
DR SMART; SM00086; PAC; 1.
DR SMART; SM00091; PAS; 2.
DR TIGRFAMs; TIGR00229; sensory_box; 2.
DR PROSITE; PS50109; HIS_KIN; 1.
DR PROSITE; PS50112; PAS; 2.
DR PROSITE; PS00245; PHYTOCHROME_1; 1.
DR PROSITE; PS50046; PHYTOCHROME_2; 1.
KW Transcription regulation; Photoreceptor; Phytochrome; Chromophore;
KW Repeat; Multigene family.
FT DOMAIN 623 694 PAS 1.
FT DOMAIN 757 828 PAS 2.
FT DOMAIN 905 1125 HISTIDINE KINASE.
FT BINDING 336 336 CHROMOPHORE (BY SIMILARITY).
FT CONFLICT 507 507 L -> S (IN REF. 2).
FT CONFLICT 586 586 L -> LQ (IN REF. 2).
SQ SEQUENCE 1132 AA; 125809 MW; 457F09C024C0F608 CRC64;

Query Match 8.3%; Score 83; DB 1; Length 1132;
Best Local Similarity 21.6%; Pred. No. 39;
Matches 43; Conservative 28; Mismatches 72; Indels 56; Gaps 8;
Oy 2 LALLEDWCRIMSVDEQKSLMTGIPADFEAEIQEVLOETLTKSLGRYLLGKIPKQENA 61
Db 586 LLILRDSFKDAEASNSKAV---VHAQLGEMELQGI--DELSVAR-----EWVRLIETA 634
Oy 62 NAVLL-----ELLEDDTDSVAIPSEVQGGVYVIFKTPNQDTEFLERLNL 108
Db 635 TAPIFAVDVEGRINGWNAKVAELTDLV--EAMGKSLVHDLVHKESQETAETKL-LFNAL 691
Oy 109 LEKGGQTVSGMFRALGQALSATV-----PCISPELLAHLIG----- 146
Db 692 RGEEDKNVEIKLRTGPEQLKAVFVWVNACSSKDYTNINIGVCFVGQDVGTQKVMDKF 751
Oy 147 -----QAMAHAPQPLP 158
Db 752 IHIQGDYKAIVHSPNPLIP 770

Search completed: September 21, 2004, 13:32:37
Job time : 6.31599 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 13:20:22 ; Search time 14.4981 Seconds
(without alignments)
694.369 Million cell updates/sec

Title: US-10-037-860-7
Perfect score: 996
Sequence: 1 PLALLEDCRIMSVDEQKSL.....EBESFEVLEQATEIVKEWP 195

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	996	100.0	195	4	US-09-189-527-7
2	443.5	44.5	462	4	US-09-189-527-13
3	441	44.3	329	4	US-09-189-527-4
4	87.5	8.8	311	2	US-08-318-837-9
5	82.5	8.3	547	4	US-09-134-000C-5974
6	81.5	8.2	580	4	US-09-328-352-7656
7	80.5	8.1	136	4	US-09-252-991A-31474
8	80.5	8.1	270	2	US-08-852-743-5
9	80.5	8.1	270	3	US-03-185-370-5
10	80.5	8.1	487	2	US-08-712-709-8
11	80.5	8.1	487	3	US-09-111-444-8
12	80.5	8.1	487	3	US-09-541-228-8
13	80	8.0	375	4	US-09-328-352-7783
14	79	7.9	258	4	US-09-252-991A-24184
15	78.5	7.9	316	1	US-08-403-634-4
16	78.5	7.9	316	1	US-08-403-634-31
17	78.5	7.9	316	3	US-08-913-441B-4
18	78.5	7.9	316	3	US-08-913-441B-31
19	78.5	7.9	445	3	US-09-457-046B-52
20	78.5	7.9	745	1	US-08-136-277-2
21	78.5	7.9	745	2	US-08-479-403-2
22	78.5	7.9	745	3	US-08-835-734-2
23	77.5	7.8	1657	1	US-08-287-959-1
24	77.5	7.8	1805	1	US-07-853-913-2
25	77	7.7	473	4	US-09-252-991A-29636
26	77	7.7	542	4	US-09-489-847-323
27	76	7.6	248	4	US-09-489-039A-12849

```

28 76 759 4 US-09-252-991A-30106 Sequence 30106, A
29 75.5 475 4 US-09-370-838-193 Sequence 193, App
30 75 419 4 US-09-134-000C-4859 Sequence 4859, Ap
31 75 1048 4 US-08-887-534A-85 Sequence 85, Appl
32 75 1048 4 US-09-527-431-85 Sequence 85, Appl
33 75 1087 4 US-09-914-259-12 Sequence 12, Appl
34 74.5 430 4 US-09-489-039A-7776 Sequence 7776, Ap
35 74.5 2108 4 US-09-252-991A-31502 Sequence 31502, A
36 74 227 4 US-09-252-991A-25961 Sequence 25961, A
37 74 388 4 US-09-489-039A-9777 Sequence 9777, Ap
38 74 407 2 US-08-934-846-4 Sequence 4, Appli
39 74 407 3 US-09-238-557-4 Sequence 4, Appli
40 74 418 2 US-08-934-846-2 Sequence 2, Appli
41 74 418 3 US-09-238-557-2 Sequence 2, Appli
42 74 521 3 US-08-948-564-14 Sequence 14, Appl
43 74 548 4 US-09-328-352-7955 Sequence 7955, Ap
44 73.5 227 2 US-08-698-407-1 Sequence 1, Appli
45 73.5 227 3 US-09-195-855-1 Sequence 1, Appli

```

ALIGNMENTS

```

RESULT 1
US-09-189-527-7
; Sequence 7, Application US/09189527A
; Patent No. 6387639
; GENERAL INFORMATION:
; APPLICANT: Jerome B. Posner
; APPLICANT: Josep O. Dalmau
; APPLICANT: Myrna R. Rosenfeld
; TITLE OF INVENTION: Ma Family Polypeptides and Anti-Ma
; TITLE OF INVENTION: Antibodies
; FILE REFERENCE: SLK98-01
; CURRENT APPLICATION NUMBER: US/09/189,527A
; CURRENT FILING DATE: 1998-11-10
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 195
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-189-527-7

```

```

Query Match Similarity 100.0%; Score 996; DB 4; Length 195;
Best Local Similarity 100.0%; Pred. No. 1.4e-98;
Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLALLEDCRIMSVDEQKSLMVTGIPADFEAEIQEVLQETLKSIGRYRLGKIFRQEN 60
Db 1 PLALLEDCRIMSVDEQKSLMVTGIPADFEAEIQEVLQETLKSIGRYRLGKIFRQEN 60
QY 61 ANAVLLELLEDTVSALPSEVQGGVWKVIFKTPNDTEFLRLNLFLEKEGOTVSGMF 120
Db 61 ANAVLLELLEDTVSALPSEVQGGVWKVIFKTPNDTEFLRLNLFLEKEGOTVSGMF 120
QY 121 RALGQEALSPTVPCISPELLAHLLGQAMAHAPOLPLPMRYRKLRFVSGSAVPAPEESF 180
Db 121 RALGQEALSPTVPCISPELLAHLLGQAMAHAPOLPLPMRYRKLRFVSGSAVPAPEESF 180
QY 181 EWLQEQATEIVKEWP 195
Db 181 EWLQEQATEIVKEWP 195

```

```

RESULT 2
US-09-189-527-13
; Sequence 13, Application US/09189527A
; Patent No. 6387639
; GENERAL INFORMATION:
; APPLICANT: Jerome B. Posner
; APPLICANT: Josep O. Dalmau
; APPLICANT: Myrna R. Rosenfeld

```

;; TITLE OF INVENTION: Ma Family Polypeptides and Anti-Ma
;; TITLE OF INVENTION: Antibodies
;; FILE REFERENCE: SLK98-01
;; CURRENT APPLICATION NUMBER: US/09/189,527A
;; CURRENT FILING DATE: 1998-11-10
;; NUMBER OF SEQ ID NOS: 14
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 13
;; LENGTH: 462
;; TYPE: PRT
;; ORGANISM: homo sapiens
US-09-189-527-13

Query Match 44.5%; Score 443.5; DB 4; Length 462;
Best Local Similarity 46.8%; Pred. No. 6e-39;
Matches 89; Conservative 34; Mismatches 64; Indels 3; Gaps 2;
QY 6 EDWCRIMSDVEQKSLMTGIPADPEAEIQEVLTSLGRYLLGKIFRKQENANAVL 65
DB 1 QWCRGHEHNRRCMLILGIPDCGEDEFEETLQEAACHLGRYRVIGRMFRRENAQAAIL 60
QY 66 LLELLEDTDVSALPSEVQGGKGVKVIKFTPNQDTEFLERLNLFLEKEGTVSGMFRALGQ 125
DB 61 LELAQDIDVALLPREIPGKGGPWEVIVKPRNSDGEFLNRLNRFLEBERTVSDMNRVLGS 120
QY 126 EALSPATVPCISPELLAHLGQMAHAPOPLL-PMRYKLRVFGSAVPAPPESEFEVWL 184
DB 121 DTNCSAPRVTISPEFWT--WAQTLGAAVQPLLEQMLYRELRFVSGNTISIPGALAFDAWL 178
QY 185 EQATEIVKWEW 194
DB 179 EHTTEMLQMW 188

RESULT 3
US-09-189-527-4
; Sequence 4, Application US/09189527A
; Patent No. 6387639
; GENERAL INFORMATION:
; APPLICANT: Jerome B. Posner
; APPLICANT: Josep O. Dalmau
; APPLICANT: Myrna R. Rosenfeld
; TITLE OF INVENTION: Ma Family Polypeptides and Anti-Ma
; FILE REFERENCE: SLK98-01
; CURRENT APPLICATION NUMBER: US/09/189,527A
; CURRENT FILING DATE: 1998-11-10
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 329
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-189-527-4
Query Match 44.3%; Score 441; DB 4; Length 329;
Best Local Similarity 46.4%; Pred. No. 6.8e-39;
Matches 90; Conservative 37; Mismatches 61; Indels 6; Gaps 4;
QY 2 LALLEDCRIMSDVEQKSLMTGIPADPEAEIQEVLTSLGRYLLGKIFRKQENA 61
DB 3 MTLLEDWCRGMDVNSQRTLLWGPVNCDEAEIETLQAAPOVS-YRMLGRMFWRENA 61
QY 62 NAVLELLEDTDVSALPSEVQGGKGVKVIKFTPNQDTEFLERLNLFLEKEGTVSGMFR 121
DB 62 KAALLELTGAVDYAAIPREMPGKGVKVIKFTPNQDTEFLERLNLFLEKEGTVSGMFR 121
QY 122 ALGQEALSPATVPCISPELLAHLGQMAHAPOPLL-PMRYKLRVFGSAVPAPPESEF 180
DB 122 VLGFQ--NPTTP--GPMPAPMUNYILDNIQPLVESIWTYKRLTLFSGKHPRAWGNF 177
QY 181 EWLLEQATEIVKWEW 194

Db 178 DPMLEHTNEVLEEW 191

RESULT 4

US-08-318-837-9
; Sequence 9, Application US/08318837
; Patent No. 5981277
; GENERAL INFORMATION:

;; APPLICANT: PRANSEN, LUCIA; DEVOS, KATHLEEN; VAN DE VOORDE,
;; APPLICANT: ANDRE; VAN HEUVERSWYN, HUGO
;; TITLE OF INVENTION: NEW POLYPEPTIDES AND PEPTIDES, NUCLEIC ACID
;; TITLE OF INVENTION: CODING FOR THEM, AND THEIR USE IN THE FIELD OF TUMOR THERAPY
;; TITLE OF INVENTION: IMMUNOLOGY
;; NUMBER OF SEQUENCES: 53
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: BIERMAN AND MUSERLIAN
;; STREET: 600 THIRD AVENUE
;; CITY: NEW YORK
;; STATE: NEW YORK
;; COUNTRY: USA
;; ZIP: 10016

;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: ASCII
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/318,837
;; FILING DATE: 13-OCT-1994
;; CLASSIFICATION: 800
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: PCT/EP 93/01022
;; FILING DATE: 28-APR-1993
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 92,401,231.3
;; FILING DATE: 30-APR-1992
;; ATTORNEY/AGENT INFORMATION:
;; NAME: CHARLES A. MUSERLIAN
;; REGISTRATION NUMBER: 19,683
;; REFERENCE/DOCKET NUMBER: 410.007
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (212) 661-8000
;; TELEFAX: (212) 661-8002
;; INFORMATION FOR SEQ ID NO: 9:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 311 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-318-837-9

Query Match 8.8%; Score 87.5; DB 2; Length 311;
Best Local Similarity 27.1%; Pred. No. 0.4;
Matches 45; Conservative 21; Mismatches 59; Indels 41; Gaps 11;
QY 53 KIFRKQENANAVL-----LLELLEDTDVSALPSEVQ---GKGWVKVIFKTPNQDTEFE 103
DB 109 KPFRDSSGANIYLEKTKELRLV-RDIRGPGVQVQCFSLQGGGLF--VEATPQDDIS--R 163
QY 104 RLNLFLKEGTVSGMFRALGQEALSPATVPCISPELLAHLG-----QAMAHAPQ 154
DB 164 RTTGFG---QVELMSGQ-RGLDLHVLSAPCRPCSTEVLLAICTSDSFVVRGFIEDVTHVPE 219
QY 155 PLLPMRY-----RKLRFVSGSAVPAPPESEFEVWLQATEIVK 192
DB 220 QQVSVIYLRVNLRLHQRKSRVQ-----PAPDSGH--WLGHVTTLLQ 259

RESULT 5
US-09-134-000C-5974
; Sequence 5974, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:

```

US-09-252-991A-31474
; Sequence 31474, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 31474
; LENGTH: 136
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31474

Query Match      8.1%; Score 80.5; DB 4; Length 136;
Best Local Similarity 21.6%; Pred. No. 0.67;
Matches 41; Conservative 24; Mismatches 48; Indels 77; Gaps 9;

QY 17 QKSLMVTGIP-----ADFEAEABIQEVLTQLKSLGRYLLGKIFRKQENANAVLLELLED 71
Db 9 QHHFRTGVKRMKLAEPFEAE-----RKLEQALALEKLKSD 46

QY 72 TDVSAIPSEVQSGKGVKVIKFTPNQDTEFLERLNLFLKEGQTVSGMFRALGQALSPA 131
Db 47 SSL-----KQEEFFKDKLQALMDKYGMTLHNIIAILDPKA--PV 83

QY 132 TVPCISPPELLAHLGQAHAPQPLLPYRKLYF-----SGSAYPAP--EESEFEVWLE 185
Db 84 TV-----SAAPO-----RRARALKYKPNNGEVVETRGNGHKKVLKAWKE 123

QY 186 Q-ATEIVKREW 194
Db 124 QYGETVESW 133

RESULT 8
US-08-852-743-5
; Sequence 5, Application US/08852743
; Patent No. 5810699
; GENERAL INFORMATION:
; APPLICANT: Force, Thomas
; APPLICANT: Kyriakis, John M.
; APPLICANT: Pombo, Celia M.
; APPLICANT: Bonventre, Joseph
; TITLE OF INVENTION: SOK-1 AND METHODS OF USE
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESS: Fish & Richardson, P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/852,743
; FILING DATE: 7-MAY-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/016,774
; FILING DATE: 7-MAY-1996
; ATTORNEY/AGENT INFORMATION:

```

```

; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 00786/327001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 270 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-852-743-5

Query Match      8.1%; Score 80.5; DB 2; Length 270;
Best Local Similarity 22.0%; Pred. No. 1.8;
Matches 44; Conservative 26; Mismatches 85; Indels 45; Gaps

Qy 2 LALLEDWCRMSVDE-----OKSLMVTGIGPADFEAEITQEVLOETLKSLGRYLRLGKIFR 56
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 73 LWIWMYEXGAGSVSDIIRLNKTL-----TDEIATILQSTLKGLEYLHFMKRIHR 123
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 57 KQENANAVLLELLEDTDVSAIPSEVQGGKGVKVIKTPNQDTEFLRLNLFLEKEGQTV 116
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 124 DIKAGNILL-----NTGSHAKLADFGVAGQLTDTMAKEN-----TV 159
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 117 SGMFRLAQGEALSPATVPCISPELLAHLGQMAHAPOP---LLPMRYRKLRVFGSAPV 173
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 160 IGTFPWMAPEVIQBIGYNCVADIWSLGTAIEMAEGRKPYADIHPMR----AIFMPTNP 215
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 174 APEESFEVWLEQATEIVKE 193
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 216 PPTFRKPKELWSDNFTDFVKQ 235
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 9
US-09-185-370-5
Sequence 5, Application US/09185370
Patent No. 6093560
GENERAL INFORMATION:
APPLICANT: Force, Thomas
APPLICANT: Kyriakis, John M.
APPLICANT: Pombo, Celia M.
APPLICANT: Bonventre, Joseph
TITLE OF INVENTION: SOK-1 AND METHODS OF USE
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESS: Fish & Richardson, P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: US
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: Fast-SEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/185,370
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/852,743
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Fraser, Janis K.
REGISTRATION NUMBER: 34,819
REFERENCE/DOCKET NUMBER: 00786/327001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906

```

;
; TELE# 200154
;
; INFORMATION FOR SEQ ID NO: 5:
;
; SEQUENCE CHARACTERISTICS:
;
; LENGTH: 270 amino acids
;
; TYPE: amino acid
;
; TOPOLOGY: linear
;
; MOLECULE TYPE: protein
;
; US-09-185-370-5
;
;
; Query Match
; Best Local Similarity 8.1%; Score 80.5; DB 3; Length 270;
; Matches 44; Conservative 26; Mismatches 85; Indels 45; Gaps 6
;
;
; QY 2 LALLEDCRIMSVDE-----QKSLMVTGIPADFFEEAEIQEVLQETLKSLGRYRLGKIFR 56
;
; Db 73 LWIWEYCGAGSVSDIIRLNKTL-----TEDEIATILQSTLKGLEYLHEWRKIHR 123
;
; QY 57 KQENANAVILLELLEDDTVSAIPSEVOQGGVKKVIKTPNQDTFELRLNLELEGQTV 116
;
; Db 124 DIKAGNTLL-----NTEGHAKLADFGVAGQITDTMAXRN-----TV 159
;
; QY 117 SCGFERALQEQALS PATVPCISPELLAHLLGQAMAHAPQ-----LLPMRYRKLRVFGSAYP 173
;
; Db 160 IGTFFWMAPEVIQIGYNCVADIWSLGLTAEIEMAEKGPYADIHPMR---AIFMPTNP 215
;
; QY 174 APEESFEVMLEQATEIVKE 193
;
; Db 216 PPTFRKPELWSDNFTDFVKQ 235
;
;
; RESULT 10
; US-08-712-709-8
; Sequence 8, Application US/08712709
; Patent No. 5863780
;
; GENERAL INFORMATION:
;
; APPLICANT: Au-Young, Janice
;
; APPLICANT: Guegler, Karl J.
;
; APPLICANT: Hawkins, Phillip R.
;
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES
;
; NUMBER OF SEQUENCES: 9
;
; CORRESPONDENCE ADDRESS:
;
; ADDRESSEE: Incyte Pharmaceuticals, Inc.

```

RESULT 10

```

US-08-712-709-8
; Sequence 8, Application US/08712709
; Patent No. 5863780
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Guegler, Karl J.
; APPLICANT: Hawkins, Phillip R.
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN K
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/712,709
; FILING DATE: Filed Herewith
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0118 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0255
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 487 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 1117791
US-08-712-709-8

```

US-08-712-709-8

Query Match 7.9%; Score 79; DB 4; Length 258;
Best Local Similarity 21.5%; Pred. No. 2.5;
Matches 41; Conservative 33; Mismatches 77; Indels 40; Gaps 8;
QY 17 QKSLMTVTGIPADFEAEIQEVLQETLKSIGRYRLGKIFRKQENANAVLL-----E 67
DB 8 RKTLLITGASSGFGQALAREALD-----AGHRVVGTV--RSEARSALAEVAPQGAQGR 59
QY 68 LIEDTDVSAIP---SEVQKGGVWKVIFKTPNQDTEFLERLNLFLKEGQTVSGMFRALG 124
DB 60 LLDVTDLAALBPTVAIERDIGPLDVNSAGYGHGILEESPLAEMRRQFEVNLF---G 116
QY 125 QEALSPATVPCISPELLAHLQMAHAPOPLPMRYRKLRFVSGSAVAPAEESFEVWL 184
DB 117 AVAMIQAVLPYMRRRRRGHILNRLHHRHLLLRQ---QVRPGRGFKP-----W- 165
QY 185 EQATEIVKEWP 195
DB 166 -----ARRWP 170
RESULT 15
US-08-403-634-4
; Sequence 4, Application US/08403634
; Patent No. 5674748
; GENERAL INFORMATION:
; APPLICANT: Giordano, Antonio
; TITLE OF INVENTION: NOVEL HUMAN CYCLIN-DEPENDENT
; TITLE OF INVENTION: KINASE-LIKE PROTEINS AND METHODS
; TITLE OF INVENTION: OF USING THE SAME
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz &
; ADDRESS: No. 5674748
; STREET: One Liberty Place, 46th floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/403,634
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/208,575
; FILING DATE: 08-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: DeLuca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: TCU-1482
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 316 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-403-634-4
Query Match 7.9%; Score 78.5; DB 1; Length 316;
Best Local Similarity 25.5%; Pred. No. 3.8;
Matches 40; Conservative 23; Mismatches 65; Indels 29; Gaps 6;
QY 5 LEDMCRIMSVDEQK-SLMVTGIPADFEAEIQEVLQETLKSIGRYRLGKIFRKQENANA 63
DB 83 LESIFLVMGYCEQDLASILENNPTFPFSEAQVKCIVLQVLRGL-QYLHRRFIHRLKVEN 141

149 DIKARNILL-----NTEGHAKLADFGVAGQLTDWAKEN-----TV 184
QY 117 SCMFALQGEALSPATVPCISPELLAHLQMAHAPO---LLPMRYRKLRFVSGSAVP 173
DB 185 IGTPWMAPEVITQIYNGYCVADISGLITAIEMAGKRPYADIHPMR-----AIFMPTNP 240
QY 174 APEESFEVWLEQATEIVKE 193
DB 241 PPTFRKPELWSDNFTDFVKQ 260
RESULT 13
US-09-328-352-7783
; Sequence 7783, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 7783
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-7783
Query Match 8.0%; Score 80; DB 4; Length 375;
Best Local Similarity 21.2%; Pred. No. 3.4; Indels 72; Gaps 10;
Matches 42; Conservative 35; Mismatches 49; Indels 72; Gaps 10;
QY 14 VDEOKSLMTVTGIPADFE---EAEIQEVL---QETLKSIGRY--RLLGKIFRKQEN 60
DB 100 IDNRQALLIATPNDPESKLNQSAPIQIVDGRNSSTAWAGSYLNKIIGQFNQKEN 159
QY 61 ANAVLLELLEDDTVSAIPSEVQKGGVWKVIFKTPNQDTEFLERLNLFLKEGQTVSGMF 120
DB 160 -----SALPISLETR--TW---YNPNQESRW----- 180
QY 121 RALQGEALSPATVPCISPELLAHLQMAHAPO-----PLIPMYRKLRFVSGSAVP 173
DB 181 -----SIMPALIAALS--MMQTLLSALSAREEQGTFDQLLVTPYTPQLIMIGRALP 232
QY 174 APEESFEVWLEQATEIV 191
DB 233 -----PIFVGLMQSTIIL 245
RESULT 14
US-09-252-991A-24184
; Sequence 24184, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24184
; LENGTH: 258
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24184


```
QY      64  VLELLEDTVSALPSEVQKGGVWKVIFK--TENQDTEFLERLNLFEKEGQTVSGMFR 121
Db      :|:  ||  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:
142 LLM-----TDKGCVKTADFGLARAYGVVFKPMTPKVVTLWYRAPELLGTTTQTTSIDMW 196
QY      122  ALGOEALSPATVPCISPELLAHLIGQAMAHAPQDPLP 158
Db      :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:  :|:
197 AVG-----CILAELLAH-----RPLLP 213
```

Search completed: September 21, 2004, 13:36:17
Job time : 15.4981 secs

Blank

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 13:20:22 ; Search time 21.3916 Seconds
(without alignments)
1272.567 Million cell updates/sec

Title: US-10-037-860-11

Perfect score: 1462

Sequence: 1 VQKGGVWKVIFKTPNQDTE.....SIEPEERDGYGRWNHEGDD 283

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_78.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	112	7.7	554	2 JMW0094	neurofilament prot
2	112	7.7	2442	2 T08621	centrosome associa
3	110.5	7.6	544	2 B44841	low molecularweig
4	105.5	7.2	659	2 S11736	resistance protein
5	105.5	7.2	659	2 S11737	resistance protein
6	103.5	7.1	991	2 C98232	hypothetical prote
7	103	7.0	374	2 A83483	ribonuclease D PAI
8	103	7.0	544	2 S07144	neurofilament trip
9	103	7.0	803	2 T39530	hypothetical prote
10	103	7.0	1840	2 T29091	transitin - chicke
11	102.5	7.0	614	2 AB2304	two-component hybr
12	102.5	7.0	879	2 AB1262	allanyl-tRNA synthe
13	101.5	6.9	520	2 AD2616	chromosomal replic
14	101.5	6.9	529	2 D97398	dnAA protein (L254
15	100.5	6.9	548	1 QFPGU	neurofilament trip
16	100.5	6.9	558	2 D70449	conserved hypothet
17	100.3	6.9	879	2 AB1625	allanyl-tRNA synthe
18	100	6.8	733	2 F82965	hypothetical prote
19	99.5	6.8	315	2 T26186	hypothetical prote
20	99.5	6.8	543	1 QFMSL	neurofilament trip
21	99.5	6.8	2712	2 T05113	hypothetical prote
22	99	6.8	399	2 G72553	RNA polymerase sig
23	98.5	6.7	1175	2 C35815	myosin heavy chain
24	98.5	6.7	1175	2 D35815	myosin heavy chain
25	98.5	6.7	1201	2 A35815	myosin heavy chain
26	98.5	6.7	1201	2 B35815	myosin heavy chain
27	98.5	6.7	1804	2 T34318	nestin - golden ha
28	98.5	6.7	2385	2 A32491	myosin heavy chain
29	98.5	6.7	2411	2 B32491	myosin heavy chain

ALIGNMENTS

RESULT 1

JMW0094

neurofilament protein-L - bovine

N:Alternate names: NF-L

C:Species: Bos primigenius taurus (cattle)

C>Date: 03-Jul-1998 #sequence_revision 10-Jul-1998 #text_change 07-May-1999

C:Accession: JMW0094

R:Hashimoto, R.; Nakamura, Y.; Goto, H.; Wada, Y.; Sakoda, S.; Kaibuchi, K.; Inagaki, M.

A:Title: Domain- and site-specific phosphorylation of bovine NF-L by Rho-associated kinase

A:Reference number: JMW0094; MUID:98238650; PMID:9571164

A:Accession: JMW0094

A:Molecule type: protein

A:Residues: 1-554 <HAS>

C:Comment: Domain- and site-specific phosphorylation by Rho-kinase regulates the assembly,

C:Superfamily: cytoskeletal keratin

Query Match 7.7%; Score 112; DB 2; Length 554;

Best Local Similarity 24.2%; Pred. No. 1.6;

Matches 55; Conservative 30; Mismatches 82; Indels 60; Gaps 10;

QY	100	FEVWLEQA---TEIVKEVPTVEAEKKWL-----AESLRG--PALD--LMHIVQADNPS	146
Db	284	FTVLTESAAKNVTDVRAAKDEVSESRLLKAKTLEIEACRGMNEALEKQLQELEDKQVAD	343
QY	147	ISVEECLEAFKQVFGSLESRRTAQVRYLKTQVE-----EGEK-----	183
Db	344	ISAMQ---DTINKLENEIRTKSEMARYLKEYQDLLNKMALDIIAAYRKLLEGETRLS	401
QY	184	-----VSAYVLRLETLRKA-----VEKRAIPRRADQVRLEQVMAGATLNQM	226
Db	402	FTSVGSLTTGTQSSQVFGSAVGLQTSSVLSARSFSPSYTSHVQEEQIEVEZETIEA-	460
QY	227	LWCLRLRLKQGGPPSPFLELMKVIREEEFEAEAFENISIEEPERDQ	273
Db	461	--AAEEAKDE--PPSEGAEEFEKEKEAEAEAEAEAEAEAEAE	503

RESULT 2

T08621

centrosome associated protein CEP250 - human

C:Species: Homo sapiens (man)

C>Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000

C:Accession: T08621

R:Mack, G.J.; Rees, J.; Sandblom, O.; Balczon, R.; Fritzler, M.J.; Rattner, J.B.

A:Title: Autoantibodies to a group of centrosomal proteins in human autoimmune sera react

A:Reference number: Z16462; MUID:98165428; PMID:9506584

A:Accession: T08621

A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: mRNA

A:Residues: 1-2442 <MAC>

adaptor protein in
pericentrin - mous
probable retroelem
hypothetical prote
conserved hypothet
neurofilament trip
hypothetical prote
heat shock protein
protein T27G7.20 [I
condensin XCAP-D2
hypothetical prote
filensin - chicken
translation elonga
protein F23F12.8 [I
hypothetical prote
myosin VII-like pr

30 98 6.7 1270 2 T09194
31 98 6.7 1320 2 A53188
32 97.5 6.7 1501 2 C84512
33 97 6.6 281 2 F75216
34 97 6.6 608 2 T03476
35 97 6.6 1087 1 QFMSH
36 97 6.6 1138 2 T24635
37 96.5 6.6 611 2 A85000
38 96.5 6.6 778 2 B86218
39 96.5 6.6 1364 2 T14900
40 96 6.6 508 2 T24622
41 96 6.6 657 2 S32739
42 96 6.6 830 2 A34347
43 96 6.6 887 2 G88484
44 95.5 6.5 300 2 A70433
45 95.5 6.5 2121 2 A59233

A;Cross-references: EMBL:AF022655; NID:g2832236; PIDN:AAC06349.1; PID:g2832237
A;Experimental source: cell line HeLa

Query Match 7.7%; Score 112; DB 2; Length 2442;
Best Local Similarity 23.4%; Pred. No. 10;
Matches 76; Conservative 47; Mismatches 94; Indels 108; Gaps 17;
QY 22 LERLNLFLKEGQTVSGMFRALGQGVSPATVPCISPELLAHLGQMAHAPOPLPM-- 79
1458 LELLSLDLKRNQVD-----LQEQIQELE-KCRS--VLEH-----LPMV 1496
Db
QY 80 --RYKLRVFGSSAVPAPEESFEVWLEQATEIVKEWPVTEAEKKRWLAESLRGPALDI- 136
1497 QEREQKLTV-QREQIREPKD-----RETQNVLEHQLLEKKQDMIESQGVQVDLK 1549
Db
QY 137 MHIVQADNPISIVE-----ECLE-AFKQVFGSLSRRTA----- 169
1550 KQVLTLECLALEENHHKWEQCKLIKELGQRETQRTVALTHLTLDLEERSQELQAQSS 1609
Db
QY 170 QVRYLKY-----QEEGKVSAYVLETLTKA-----VVKR----- 202
1610 QIHDLESHSTVLARELQEQDQVKSQREQIELOQKEHLTQDLRRDQELMLQKERIOV 1669
QY 203 -----AIPRIADQVRLEQVMAG---ATLNQMLWCLRELKDDQGGPPPS-----FILE 245
1670 LEDQRTQTKILEEDLEQKLSRERGRLTTQRL-----MQEAEKGKPSKAQSGSL 1725
Db
QY 246 LMKVIREEEERASFNESIEPEE 270
1726 HMKLIIRDKEVECCQEHIEHQE 1750
Db

RESULT 3
B44841
low molecular weight neurofilament protein XNF-L - African clawed frog
C;Species: Xenopus laevis (African clawed frog)
C;Date: 31-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 12-Apr-1995
C;Accession: B44841
R;Chatnas, L.R.; Szaro, B.G.; Gainer, H.
J. Neurosci. 12, 3010-3024, 1992
A;Title: Identification and developmental expression of a novel low molecular weight neu
A;Reference number: A44841; MUID:92356194; PMID:1494944
A;Accession: B44841
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-544 <CHA>
A;Experimental source: brain
A;Note: sequence inconsistent with the nucleotide translation
A;Note: sequence extracted from NCBI backbone (NCBIN:110225, NCBIP:110226)
C;Superfamily: cytoskeletal keratin

Query Match 7.6%; Score 110.5; DB 2; Length 544;
Best Local Similarity 22.3%; Pred. No. 2;
Matches 59; Conservative 30; Mismatches 83; Indels 93; Gaps 10;
QY 80 RYKLRVFGSSAVPAPEESFEVWLE-----QATEIVKEWPVTEAEKKRWLA-- 126
258 QYEKL-----AAKMQSAEEFKSRFTVLTQSAARNTDAVRAAKDENESERMLSAK 309
QY 127 -----ESLRGPALDLMHIVQ--ADNPISIVBECLEAFKQVFGSLSRRTAQVRYLKYQE- 179
310 GLEIACRGVNEALQRIQIELEDKQSGEIQAGMQDAINKLEELNTRKSEMARYLKEYQDL 369
QY 180 -----BEKV-----SAYVLRLETLRKA 198
370 LNVKMLDIEIAAYRKLEGETRSLFSGVGAITSYGTQSAVFGRSAYLSQSSYM--- 426
QY 199 VEKAIIPRIADQVRLEQVMAGATLNQMLWCLRELKDDQGGPPPSFLMLKVIREEESEA 258
427 -TSRAFFPYGSHVQEQGLDIETIES---SRAEAKAEAP-----EEEEEA 471
QY 259 SFENESIEPEERDGYGWNHEGDD 283

Db 472 EEEGGGSEAESE-----GEEGEE 490
RESULT 4
S11736
resistance protein Mx2, interferon-regulated - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 19-May-2000
C;Accession: S11736
R;Meier, E.; Kunz, G.; Haller, O.; Arnheiter, H.
submitted to the EMBL Data Library, April 1990
A;Description: Unexpected activity of rat Mx proteins against a Rhabdovirus.
A;Reference number: S11735
A;Accession: S11736
A;Molecule type: mRNA
A;Residues: 1-659 <EMB>
A;Cross-references: EMBL:X52712; NID:g56722; PIDN:CAA36936.1; PID:g56723
C;Superfamily: dynamin-related protein VPS1

Query Match 7.2%; Score 105.5; DB 2; Length 659;
Best Local Similarity 21.5%; Pred. No. 5.8;
Matches 64; Conservative 51; Mismatches 84; Indels 99; Gaps 16;
QY 20 EFLERLNLFLKEGQTVSGMFRALGQGVSPATVPCISP-----ELLAHLGQMAHAPO 75
292 EALQKEQVFFKEHPQ-----FRALLEDG--KATVPCLAERLTMELISHICKS----- 336
Db
QY 76 LLPMYRYKLRVFGSSAVPAPEESFEVWLEQATEIVKEWPVTEAEKKRWLAESLRGPALD 135
337 -LPLENQIK-----ESHQSTSEELQKYGADIPEDENEXTLFLIEKINAFNQD 383
QY 136 LMHIVQADNPISIVBEC---LEAFKQVF-GSLESRTAQ-----VRYLKY-----QEEGK 183
384 ITAIVEGEE-IVREKECRLFTKLKKEFFLWSEIERNFQKSGDALYKEVYTFMFQYRGRE 442
QY 184 VSAVY-----LRLLETLRKAVE-----KRAIPR 206
443 LQGFVNYKTFENIRROIKTLEBPAMEMLHKVTEIVRAAFTTVSEKNFSEFFNLHRTKS 502
QY 207 RIADQVRLEQ-----VMAGATLNQMLWCLRELKDDQGGPPPSFLMLKVIREEESEE 257
503 KLED-IRLEQTEAEAKSIRLHFQMEQIIVC-----QDQ-----IVRKALQKVREEEAEE 550
Db

RESULT 5
S11737
resistance protein Mx3, interferon-regulated - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 17-Mar-2000
C;Accession: S11737
R;Meier, E.; Kunz, G.; Haller, O.; Arnheiter, H.
submitted to the EMBL Data Library, April 1990
A;Description: Unexpected activity of rat Mx proteins against a Rhabdovirus.
A;Reference number: S11735
A;Accession: S11737
A;Molecule type: mRNA
A;Residues: 1-659 <EMB>
A;Cross-references: EMBL:X52713; NID:g56724; PIDN:CAA36937.1; PID:g56725
C;Superfamily: dynamin-related protein VPS1

Query Match 7.2%; Score 105.5; DB 2; Length 659;
Best Local Similarity 21.5%; Pred. No. 5.8;
Matches 64; Conservative 51; Mismatches 84; Indels 99; Gaps 16;
QY 20 EFLERLNLFLKEGQTVSGMFRALGQGVSPATVPCISP-----ELLAHLGQMAHAPO 75
292 EALQKEQVFFKEHPQ-----FRALLEDG--KATVPCLAERLTMELISHICKS----- 336
Db
QY 76 LLPMYRYKLRVFGSSAVPAPEESFEVWLEQATEIVKEWPVTEAEKKRWLAESLRGPALD 135
337 -LPLENQIK-----ESHQSTSEELQKYGADIPEDENEXTLFLIEKINAFNQD 383
QY 136 LMHIVQADNPISIVBEC---LEAFKQVF-GSLESRTAQ-----VRYLKY-----QEEGK 183

Db 384 ITAIVEGEE-IVREKECLFTLRKEFFLWSEIERNFQKGDALYKEVYTFEMQVGRG 442
 QY 184 VSAYV-----LRLETLRKAVE-----KRAIPR 206
 Db 443 LPGFVNYKTFENIRROQKLTLEPAMVLMHKTVEIVRAAFTTVSEKNFSEFNLRHTTKS 502
 QY 207 RIADQVRLEQ-----VMAGATLQMLWCRLRLKDKQPPSPFLMKVIREEEEE 257
 Db 503 KLED-IRLEQETEAEKAIHLHFQMQIIYC-----QDQ-----IYRKALQKVREBAEE 550

RESULT 6

C98232
 Hypothetical protein AGR_L1617 [imported] - Agrobacterium tumefaciens (strain C58, Cere
 C;Species: Agrobacterium tumefaciens
 C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 18-Nov-2002
 C;Accession: C98232

R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorllo, B.; Goldman,
 A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
 Science 294, 2323-2328, 2001
 A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
 A;Reference number: A97359; MUID:21608551; PMID:11743194

A;Accession: C98232
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-991 <KUR>
 A;Cross-references: GB:AE007870; PIDN:AAK99381.1; PID:g15159232; GSPDB:GN00170
 C;Genetics:
 A;Gene: AGR_L1617
 A;Map position: linear chromosome

Query Match 7.1%; Score 103.5; DB 2; Length 991;
 Best Local Similarity 24.6%; Pred. No. 14;
 Matches 75; Conservative 34; Mismatches 101; Indels 95; Gaps 15;

QY 41 RALQOEGVSPATVPCISPE-----LLAHLGQAMAHAPQPLPMRYKLRVFSASV 92
 Db 81 RALG-----GTVPGIWPEIWMNSRIEAGFRGEVNAVYRDQPTLMRH----- 123
 QY 93 PAPEEESFEVWLEQATVKE-----WVTEAEKRWLAESIRGPDALDMHIVQA-- 142
 Db 124 GVPBEVDFDY-----TPIYEGGTVGVCLTVENTDKVRALEAQSMELSLTNALP 180
 QY 143 -----DNPSI---SVBCELEAF---KQVPGS-----LESRTAQVRLKYQE 179
 Db 181 ILVGYVDVRYVPFANDGYLEWFGRAEVLGRSVPDIVGNAFPAARTYLDRAI----- 235
 QY 180 EGEK-USAYVLRLETLRKAVEKRAIPRIAD-----QVRLEQVM----- 218
 Db 236 AGEKIVSDTVIRPDGSLRAAEISYVPRTISDGSIDGIYVLIIDIEERKSEQEILISNN 295
 QY 219 ----AGATLQMLWCRLRLKDKQPPSPFLMKVIREEEEEASFNESIEPEERDQ-- 273
 Db 296 RFAAEEVAGVGLWNTSADGMRGEQPAWTAM---TGQTPPEYQDFGWDADVHPEDRQGS 352
 QY 274 YGRWN 278
 Db 353 VDSWN 357

RESULT 7

A83483
 ribonuclease D PA1294 [imported] - Pseudomonas aeruginosa (strain PA01)
 C;Species: Pseudomonas aeruginosa
 C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
 C;Accession: A83483

R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
 ; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
 A;Reference number: A82950; MUID:20437337; PMID:10984043

A;Accession: A83483
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-374 <STO>
 A;Cross-references: GB:AE004559; GB:AE004091; NID:g9947228; PIDN:AAG04683.1; GSPDB:GN001
 A;Experimental source: strain PA01
 C;Genetics:
 A;Gene: rnd, PA1294
 C;Superfamily: ribonuclease D

Query Match

Best Local Similarity 7.0%; Score 103; DB 2; Length 374;
 Matches 65; Conservative 34; Mismatches 90; Indels 92; Gaps 14;

QY 23 ERLNIFLKEGQTVSGMFRALQEGVSPATVPCISPELLAHLGQAMAHAPQPLPMRYR 82
 Db 87 EDLEVLRLTGS-----LPVLPDTQLAAAVLG--MAHS-----MGVS 122
 QY 83 KLRVFGSAVPAPPEEESFEVWLEQATEIVKSWPVTAEAKRWLABSLRGPALDMHIVQA 142
 Db 123 KL-VKEVLIDILPKDTRSDLQR-----PLTEMQ-----MRYAADDVQHIAQV 165
 QY 143 -----DNPSISVEEC-----LEAPKQV-FG-SLESRTAQVRYLAKTY 177
 Db 166 YLALDTRLSEKRAWLLEDGAEIVANLCRESDPREAYREVKLGWLRPQQLAVRELCAW 225
 QY 178 QEGEKV-----SAYVLRLETLRKAVEKRAIPRIADQVRLEQ-----V 217
 Db 226 REEQARLRNPRNHVLRERTLWPLA--RLLPKNKTDLAAIEDMHPTVRQDGDPLIELI 282
 QY 218 MAGATLQMLWCRLRLKDKQPPSPFLMKVIREEEEEEA 258
 Db 283 AQAARLPQSEW---PEALPELPPEVTPLLKSLRAIGQREA 320

RESULT 8

S07144

neurofilament triplet L protein - human

N;Alternate names: neurofilament light polypeptide (68k)

N;Contains: Glu-50 brain peptide

C;Species: Homo sapiens (man)

C;Date: 29-Jan-1993 #sequence_revision 29-Jan-1993 #text_change 21-Jul-2000

C;Accession: S07144; I52832; A60703

R;Julien, J.P.; Grosvelde, F.; Fazdanbakhsh, K.; Flavell, D.; Meijer, D.; Mushynski, W.

Biochim. Biophys. Acta 909, 10-20, 1987

A;Title: The structure of a human neurofilament gene (NF-L): a unique exon-intron organi:

A;Reference number: S07144; MUID:87214213; PMID:3034332

A;Accession: S07144

A;Molecule type: DNA

A;Residues: 1-544 <JUL>

A;Cross-references: EMBL:X05608; NID:g1495072; PIDN:CAA29097.1; PID:g1279504

R;Pospelov, V.A.; Pospelova, T.V.; Julien, J.P.

Cell Growth Differ. 5, 187-196, 1994

A;Title: AP-1 and Krox-24 transcription factors activate the neurofilament light gene pr

A;Reference number: I52832; MUID:94235564; PMID:8180132

A;Accession: I52832

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-10 <POS>

A;Cross-references: GB:S70309; NID:g547176; PIDN:AAD14057.1; PID:g4261757

R;Nomata, Y.; Watanabe, T.; Wada, H.

J. Biochem. 93, 825-831, 1983

A;Title: Highly acidic proteins from human brain: purification and properties of Glu-50

A;Reference number: A60703; MUID:83265667; PMID:6135695

A;Accession: A60703

A;Molecule type: protein

A;Residues: 469-472; D, 474 <NOM>

A;Experimental source: Glu-50 brain peptide

A;Note: this acidic protein is named for its greater than fifty per cent glutamic acid c

C;Genetics:

A;Gene: GDB:NEFL; NFI

A;Cross-references: GDB:120227; OMIM:162280

A;Map position: 8p21-8p21
A;Introns: 349/3; 391/2; 498/1
C;Superfamily: cytoskeletal keratin
C;Keywords: brain; coiled coil; intermediate filament
P;469-544/Product: Glu-50 peptide #status predicted <E50>

Query Match 7.0%; Score 103; DB 2; Length 544;
Best Local Similarity 24.2%; Pred. No. 6.9; Mismatches 87; Indels 58; Gaps 12;
Matches 59; Conservative 30;

QY 100 FEVWLEQA---TEIVKWPVTAERKRWL-----AESLRG--PALD--LMHIVQADNPS 146
DB 285 FTVLTSAKNTDAVRAAKDEVSERULLKAKTLEIACRGWNALEKQLQLELDKQAD 344
QY 147 ISVEECLFAKQVFGSLSSRRTAQVRYLKYQE-----EKEK---- 183
DB 345 ISAMQ--DTINKLENLRTTKSEMARYLYKEYQOLLNVKMDIEIAAYRKLLGESETRLS 402
QY 184 -----VSAYVLRLETLRLKA-----VEKRAIPRRIADQVRLQVNMAGATLNM 226
DB 403 FTSVGSITSGYSQSQVFGRSAYGGLQTSYLMSTRSPSYTYTSHVQEQTEVEETIEA- 461
QY 227 LWCRLRELKQDQPPSPFLFMKVIREE-EEERASPENESIER-----PEERDGYGRWNH 279
DB 462 --SKAEAKDEPPSGEAEERKDEEAEAEAEAEAEAEAEAEAEAEAEAEAEAEAEAEAE 516
QY 280 EGGD 283
DB 517 EGEE 520

RESULT 9
T39530
hypothetical protein SPBC1685.14c - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C;Accession: T39530
R;Wood, V.; Rajandream, M.A.; Barrell, B.G.; Hilbert, H.; Duesterhoeft, A.
submitted to the EMBL Data Library, July 1998
A;Reference number: Z21861
A;Accession: T39530
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-803 <WOO>
A;Cross-references: EMBL:AL011154; PIDN:CAA20062.1; GSPDB:GNO0067; SPDB:SPBC1685.14c
A;Experimental source: strain 972h-; cosmid c1685
C;Genetics:
A;Map position: 2
A;Gene: SPBC1685.14c
A;Introns: 8/2; 41/2

Query Match 7.0%; Score 103; DB 2; Length 803;
Best Local Similarity 20.1%; Pred. No. 11; Mismatches 89; Indels 138; Gaps 15;
Matches 67; Conservative 40;

QY 15 PNQDTEFLRLNLFLEKEGQTVSGMPRALQEGVSPATVPCISPELLAHLLGQMAHAPQ 74
DB 169 PSKEVDSLENIVTKLDLSE---DIMLKKQQLDDEIAK-----KYLQ--QEAEE 216
QY 75 PLLPMRYKLRVFGSGAVPAPEE-----SFEVWLEQATEIV---- 111
DB 217 PLVQQQ-----TSIVNPEKEVTKENIKSLEGELMTGISAEHLFDVAVEEPIQ 267
QY 112 -----KEPVTVEAKRWLAELRGPALDLMHIVQAD--NPSISVEECLEAF 156
DB 268 DPNVEASVFDLGDWNYFTTISTEKTWLSQS-----VDADWNPVFSFE----- 310
QY 157 KQVFGSLESRTAQVRYLKYQERGEKVSAYVLRLETLRLKAVEKRAIPRRIADQVRLQ 216
DB 311 -----HLSTFWNVFDANSNAFSLWLPDS-----QVRMEQ 340
QY 217 VM-----AGATLQMLCRL-----RELKQDQPPSPFLFMKVIREEE 255

DB 341 FQELLMRALWESLNQQRWLKIDDRQDYVMTTFHEDELED-SEDEEPARQOLLRSKEEE 399
QY 256 EEA-----SPENESIEEPERDGYGRWNHE 280
DB 400 EEEDEASDFEDSPADFDGEGADDLD-ESRWKE 432

RESULT 10
T29091
transitin - chicken
C;Species: Gallus gallus (chicken)
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T29091
R;Cole, G.J.
submitted to the EMBL Data Library, April 1996
A;Reference number: Z20560
A;Accession: T29091
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-1840 <COL>
A;Cross-references: EMBL:X80877; NID:g1279430; PID:e237634; PIDN:CAA56845.1

Query Match 7.0%; Score 103; DB 2; Length 1840;
Best Local Similarity 25.3%; Pred. No. 32; Mismatches 87; Indels 82; Gaps 14;
Matches 68; Conservative 32;

QY 20 EFLERLNLFLKEGQTVS-----GMPRALQEGVSPATVPCISPELLAHLLGQA- 68
DB 35 KFLSENEGIRAEIOSTKENPAGTHAGPGMRSSCGRSG-----CAAPRLHREVCGRAG 87
QY 69 -----MAHAP---QPLLPMRYKLRVFGSGAVPAPEESFEVWL-EOATEIVKEWPT 117
DB 88 RDNLYBEVQVHRSRCQKEQAEEAKRQLSSSKKELEERRAQIWLKERAVQLEKE--V 144
QY 118 EAEKKRWLAESLRGPAIDLMHIVQADNPSISVEECLEAFKQVFGSLESRTAQVRYLKY 177
DB 145 EA-----LLEVHEEKAGLDQE--LASFSQ---SLEGRFCAVPVF---- 179
QY 178 QBEQEKVSAYVLRLETLRLKAVEKRAIPRRIADQVRLQVNMAGATLQMLW----- 228
DB 180 --QPVEVEDYSKRLSEIWRGAVETVK---AEVSQLERALGQAKEN--LWQVAEDNQQS 230
QY 229 -CRLRELKQDQPPSPFLFMKVIREEE 256
DB 231 QLQLRHLKE-----LVGLKVRKEMLEE 253

RESULT 11
AB2304
two-component hybrid sensor and regulator all3985 [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C;Accession: AB2304
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriiguchi, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AB2304
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-614 <KUR>
A;Cross-references: GB:BA000019; PIDN:BAH75684.1; PID:g17133120; GSPDB:GNO0179
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: all3985

Query Match 7.0%; Score 102.5; DB 2; Length 614;
Best Local Similarity 21.5%; Pred. No. 8.7; Mismatches 89; Indels 81; Gaps 9;
Matches 56; Conservative 34;

QY 76 LLPMYRKLRFVSGSAV--PAPEESFEVWLEQATEI-----VKE 113

Db 48 VLPVRNDGCEIFAGMTITQETTERKQSELAQNLQLELHIDTTPLAVVQWQDCHLCVTR 107
 Qy 114 WPVTEAEKKRWLAESLRGALDLMHIVQADNPISVECELEAFKQVFGSLSRRTAQ-VR 172
 Db 108 WSSAAEKIFGLAEVIGIKYIQDLHIV-----YEDMTAVAFVSRLSLSGQEAIIQ 159
 Qy 173 YLKIT-----QREGKVSAYVLRLETLRLKAVBK----- 201
 Db 160 YNRNYTKDGRVYCEWYNSITNEGSVTSVLVDVTERIQAEKALRQSELWTERA 219
 Qy 202 RAIPIRIADQVRLEQVM-----AGATINQML-WCRLELKDQGGPPPSFLELM 247
 Db 220 QALERERAAARIELERASRMKDEFLAIVSHELSPILNGILGWSRLRLTRKLSF-----EKI 274
 Qy 248 KVIREEEEEEFENESTEE 267
 Db 275 EQALESIBRNAQAQTOLIED 294

RESULT 12

alanyl-tRNA synthetase [imported] - Listeria monocytogenes (strain EGD-e)
 C:Species: Listeria monocytogenes
 C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 17-Mar-2003
 C:Accession: AH1262
 R:Glasier, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.
 Science 294, 849-852, 2001
 A:Authors: Kreit, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Mak, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of Listeria species.
 A:Reference number: AB1077; MUID:21537279; PMID:11679669
 A:Accession: AH1262
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-879 <GLA>
 A:Cross-references: GB:NC 003210; PIDN:CAC99582.1; PID:gl6410933; GSPDB:GN00177
 A:Experimental source: strain EGD-e
 C:Genetics:
 A:Gene: alas
 C:Superfamily: alanyl-tRNA ligase

Query Match 7.0%; Score 102.5; DB 2; Length 879;
 Best Local Similarity 24.1%; Pred. No. 14;
 Matches 70; Conservative 37; Mismatches 110; Indels 73; Gaps 16;
 Qy 13 KTPNQDTFELRLNLFLEKEGQTVSG-----MPRALGQEGVSPATVPCISPELLAHLG 66
 Db 527 KAPNKQN--LHRISV--KEGVLTGDTVKLAVDKVRRETIKNHTATHLLHRAKLDTLG 581
 Qy 67 QAMAHAPQPLPWRKLRVFGSAVPAPEESF-----BWLQEAIVKEWPVT 117
 Db 582 EHVNAQSLVSPDRLR----FDFSHFQGITEEELTKMEEIVNEKIW-EQINVVIEEMPIA 636
 Qy 118 EAEKKRWLAESLRGAL-DLMHIVQADNPISVECELEAFKQVFGSLSRRTAQVRYLKT 176
 Db 637 EA--KELGAMALFGEKYGDIVRVQVG--KYSIELC-----GGVHVNTADIGLFKI 684
 Qy 177 YQEGEKVSAYVLRLETLRLKAVEKRAIPRIADQVRLEQVMAGATINQMLWCRLELKD 236
 Db 685 VSETG--IGAGTRRIE-----AVTGKAYRFPVTQEN-----TLKQAASLLKTTTKE 729
 Qy 237 QGPPPSFLELMKVIREEEAEFENESI-----EPBERDG 273
 Db 730 ---TPQKVELL---QADLRVREKRENSLLSKLASAASADIFESPBIGG 772

RESULT 13

AD2616
 Chromosomal replication initiator protein DnaA [imported] - Agrobacterium tumefaciens (s
 C:Species: Agrobacterium tumefaciens

C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
 C:Accession: AD2616
 R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.; Erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClellan, Karp, P.; Romero, P.; Zhang, S.
 Science 294, 2317-2323, 2001
 A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, J. ster, E.W.
 A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
 A:Reference number: AB2577; MUID:21608550; PMID:11743193
 A:Accession: AD2616
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-520 <KUR>
 A:Cross-references: GB:AE008688; PIDN:AAL41346.1; PID:gl7738660; GSPDB:GN00186
 A:Experimental source: strain C58 (Dupont)
 C:Genetics:
 A:Gene: dnaA
 A:Map position: circular chromosome
 C:Superfamily: replication initiation protein dnaA

Query Match 6.9%; Score 101.5; DB 2; Length 520;
 Best Local Similarity 26.4%; Pred. No. 8.3;
 Matches 51; Conservative 30; Mismatches 83; Indels 29; Gaps 8;
 Qy 20 EFLERLNLFLEKEGQTVSGMFRALGQEGVSPATVPCISPELLAHLGQAMAHAPQPLLP 79
 Db 294 EFCHELLNMLLDSAKQVVVAADRA-----PWELESLDPRVRSRLOQGMATIEGPDYDM 346
 Qy 80 RYRKLRFVSGSAVPAPEESFEVWLEQATEIVKEWPTEAEKKRWLAESLRGALDLMHI 139
 Db 347 RYEMLNRMGSA--RQDDPSFEISDEILTHVAKS--VTASGRE-----LEG-AFNQLMP 395
 Qy 140 VQADNPISVECELEAFKQVFGSLSRRTAQVRYLKTQYQEGEKVSA---YVLELTLLR 196
 Db 396 RRSFEPNLSVDRVDELLSHLVSGEAKR---VRI-----EDIQIRVARHYNVSQELVSN 447
 Qy 197 KAVEKRAIPRIA 209
 Db 448 RRTRVIVKPRQIA 460

RESULT 14

D97398
 dnaA protein (I25439) [imported] - Agrobacterium tumefaciens (strain C58, Cereon)
 C:Species: Agrobacterium tumefaciens
 C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 18-Nov-2002
 C:Accession: D97398
 R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman, A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.; Science 294, 2323-2328, 2001
 A:Title: Genome sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens C58.
 A:Reference number: A97359; MUID:21608551; PMID:11743194
 A:Accession: D97398
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-529 <KUR>
 A:Cross-references: GB:AE007869; PIDN:AAK86141.1; PID:gl5155230; GSPDB:GN00169
 C:Genetics:
 A:Gene: AGR_C.566
 A:Map position: circular chromosome
 C:Superfamily: replication initiation protein dnaA

Query Match 6.9%; Score 101.5; DB 2; Length 529;
 Best Local Similarity 26.4%; Pred. No. 8.5;
 Matches 51; Conservative 30; Mismatches 83; Indels 29; Gaps 8;
 Qy 20 EFLERLNLFLEKEGQTVSGMFRALGQEGVSPATVPCISPELLAHLGQAMAHAPQPLLP 79
 Db 303 EFCHELLNMLLDSAKQVVVAADRA-----PWELESLDPRVRSRLOQGMATIEGPDYDM 355
 Qy 80 RYRKLRFVSGSAVPAPEESFEVWLEQATEIVKEWPTEAEKKRWLAESLRGALDLMHI 139

Db 356 RYEMLNRRMGSA--RQDDPSFEISDEILTHVAKS--VTASGRE-----LEG-AFNQLMF 404
 QY 140 VQADNPSSVVECLAEAFKQVFGSLRSRTAQVRYLKTQESEKVS-----YVLRLETLR 196
 Db 405 RRSPEPNLSVDRVDLLSHLVGSEAKR---VRI-----EDIQRIARHYNVSRDELVS 456
 QY 197 KAVEKRAIPRIA 209
 Db 457 RTRVIVPRQIA 469

RESULT 15

QPPGL
 neurofilament triplet L protein - pig
 N:Alternate names: 68K neurofilament protein
 C:Species: Sus scrofa domestica (domestic pig)
 C>Date: 15-Nov-1984 #sequence revision 28-May-1986 #text change 10-Oct-1997
 C:Accession: A91337; A90973; A34569; A02963
 R:Geisler, N.; Plessmann, U.; Weber, K.
 PDBS Lett. 182, 475-478, 1985
 A:Title: The complete amino acid sequence of the major mammalian neurofilament protein
 A:Reference number: A91337; MUID:85f54583; PMID:3920075
 A:Accession: A91337
 A:Molecule type: protein
 A:Residues: 1-547 <GE1>
 R:Geisler, N.; Kaufmann, B.; Fischer, S.; Plessmann, U.; Weber, K.
 EMBO J. 2, 1295-1302, 1983
 A:Title: Neurofilament architecture combines structural principles of intermediate filament
 A:Reference number: A90973
 A:Accession: A90973
 A:Molecule type: protein
 A:Residues: 1-82; 278-548 <GE2>
 A>Note: Residue 322 is either lysine or arginine
 R:Conda, Y.; Nishizawa, K.; Ando, S.; Kitamura, S.; Minoura, Y.; Nishi, Y.; Inagaki, M.
 Biochem. Biophys. Res. Commun. 167, 1316-1325, 1990
 A:Title: Involvement of protein kinase C in the regulation of assembly-disassembly of ne
 A:Reference number: A34569; MUID:90211318; PMID:2108674
 A:Accession: A34569
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 9-14; 23-29; 30-53 <GON>
 C:Comment: Mammalian neurofilaments usually contain three polypeptides, L, M, and H (wit
 ke all other intermediate filament proteins: a conserved alpha-helical region, whose hel
 C:Comment: The amino-terminal headpiece is basic with a high content of hydroxyamino aci
 al beta turns; domain b is acidic and rich in glutamic acid and lysine residues.
 C:Comment: The extra mass and high charge density that distinguish the neurofilament pro
 charged scaffolding structure suitable for interaction with other neuronal components
 C:Comment: The boundaries of the domains between residues 70-92 and 399-402 are not yet
 C:Comment: This protein was isolated from spinal cord.
 C:Superfamily: cytoskeletal keratin
 C:Keywords: coiled coil; intermediate filament
 F:1-70/Domain: head <HED>
 F:92-123/Domain: coil 1a, alpha-helical rod <RIA>
 F:137-232/Domain: coil 1b, alpha-helical rod <R1B>
 F:255-399/Domain: coil 2, alpha-helical rod <RD2>
 F:402-548/Domain: tail <TAI>
 F:402-442/Region: tail subdomain a
 F:443-548/Region: tail subdomain b

Query Match 6.9%; Score 100.5; DB 1; Length 548;
 Best Local Similarity 19.9%; Pred. No. 11;
 Matches 56; Conservative 36; Mismatches 90; Indels 99; Gaps 11;
 QY 80 RYKLRVFGSAVPAPEESFEVWLEQATEIVKEWPTVTAEKRWLAESLRGPDALDLMHI 139
 Db 263 QYEKL-----AAKMQNAEEWFKSRFTVLTE-----SAAKNTDAVRAAKDEVSESRRL 310
 QY 140 VQADNPSSVVECL---EAFKQVFGSLRSRTAQV-----RYL 174
 Db 311 LKA--KTLIEACXGWNLEAKQLQLEDKQNADISAMQDTINKLENLRTTKSEMARYL 368
 QY 175 KTYQE-----EGEK-----VSAYVLELTLRKA----- 198

Db 369 KEYQDLNKKVALDIEIAAYRKLLEGEETRLSFTSVGSLTTGYSSQVFGRSAYGGLOT 428
 QY 199 -----VEKRAIPRIADQVRLEQVWAGATLNQMLWCRLELKDQPPSPFLELMKVIRE-- 252
 Db 429 SSYLMSTRSPFSYTYTSHVQEQIEVEETIEA---AKAEAKDE--PPSEGEAEBSGKEKE 483
 QY 253 -----EEEEASPENBSIEEPERDGYGRWNHE 280
 Db 484 EAAEAAEAEEGAQEEEAKEBSEAKBEGGEGEQGEE 524

Search completed: September 21, 2004, 13:40:42
 Job time : 24.3916 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 13:20:21 ; Search time 7.71499 Seconds
(without alignments)
1910.028 Million cell updates/sec

Title: US-10-037-860-11

Perfect score: 1462

Sequence: 1 VQGGGVKVKVIFKPNQDTE.....SIEEPEERDGYGRWNHEGDD 283

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1459	99.8	364	1 PMA2_HUMAN	Q9UL42 homo sapien
2	1428	97.7	364	1 PMA2_MACFA	Q9Gmu3 macaca fasc
3	1193.5	81.6	365	1 PMA2_MOUSE	Q8bhk0 mus musculus
4	634	43.4	353	1 PMA1_RAT	Q8vHz4 rattus norv
5	628	43.0	353	1 PMA1_HUMAN	Q8nd90 homo sapien
6	617	42.2	353	1 PMA1_MOUSE	Q8clc8 mus musculus
7	596.5	40.8	351	1 MOP1_HUMAN	Q96by2 homo sapien
8	594	40.6	351	1 MOP1_MACFA	Q95k14 macaca fasc
9	571.5	39.1	352	1 MOP1_MOUSE	Q9erh6 mus musculus
10	113	7.7	2442	1 CEP2_HUMAN	Q9bv73 homo sapien
11	112	7.7	554	1 NFL_BOVIN	Q02548 bos taurus
12	110.5	7.6	544	1 NFL_XENLA	P35616 xenopus lae
13	110.5	7.6	1714	1 ITN1_MOUSE	Q920r4 mus musculus
14	108.5	7.4	1217	1 ITN1_RAT	Q9wve9 rattus norv
15	107	7.3	525	1 NAB2_MOUSE	Q61127 mus musculus
16	107	7.3	1721	1 ITN1_HUMAN	Q15811 homo sapien
17	106	7.3	541	1 NFL_RAT	P19527 rattus norv
18	105.5	7.2	659	1 MX2_RAT	P18589 rattus norv
19	105.5	7.2	659	1 MX3_RAT	P18590 rattus norv
20	105	7.2	582	1 PESC_BRARE	P79741 brachydanio
21	105	7.2	759	1 CHLD_ARATH	Q9ejel arabidopsis
22	104.5	7.1	525	1 NAB2_HUMAN	Q15742 homo sapien
23	103.5	7.1	548	1 NFL_PIG	P02547 sus scrofa
24	103	7.0	543	1 NFL_HUMAN	P07196 homo sapien
25	102.5	7.0	879	1 SVA_LISMO	Q9722 listeria mo
26	101.5	6.9	487	1 DNAA_AGRFS	Q8u1hi agrobacteri
27	101.5	6.9	1330	1 KTN1_VULVU	Q97961 vulpes vulp
28	100.5	6.9	558	1 YH32_AQUAE	O67622 aquifex aeo
29	100.5	6.9	879	1 SVA_LISIN	Q92bk9 listeria in
30	99.5	6.8	542	1 NFL_MOUSE	P08551 mus musculus
31	99	6.8	399	1 RPSD_THEMA	P77994 thermotoga
32	98.5	6.7	1962	1 MYSA_DROME	P05661 drosophila
33	98	6.7	1172	1 CNA2_MOUSE	Q90xc6 mus musculus

RESULT 1

ID	PMA2_HUMAN	STANDARD	PRT	364 AA
AC	Q9UL42; Q94959; Q95145; Q9UL43;			
DT	15-MAR-2004 (Rel. 43, Created)			
DT	15-MAR-2004 (Rel. 43, Last sequence update)			
DT	15-MAR-2004 (Rel. 43, Last annotation update)			
DE	Paraneoplastic antigen Ma2 (Onconeural antigen MA2) (Paraneoplastic neuronal antigen MM2) (40 kDa neuronal protein).			
GN	PMA2 OR MA2 OR KIAA0883.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Euthera; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]_TaxID=9606;			
RP	SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.			
RC	TISSUE=Cerebellum;			
RX	MEDLINE=99270611; PubMed=10362822;			
RA	Voltz R., Gultekin S.H., Rosenfeld M.R., Gerstner E., Eichen J.,			
RA	Posner J.B., Dalmau J.;			
RT	"A serologic marker of paraneoplastic limbic and brain-stem encephalitis in patients with testicular cancer.";			
RT	New Engl. J. Med. 340:1788-1795(1999).			
RL	[2]			
RN	SEQUENCE FROM N.A.			
RP	TISSUE=Brain;			
RC	MEDLINE=99156230; PubMed=10048485;			
RX	Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirosewa M.,			
RA	Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;			
RT	"Prediction of the coding sequences of unidentified human genes. XII. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro.";			
RT	DNA Res. 5:355-364(1998).			
RL	[3]			
RN	IDENTIFICATION, AND SUBCELLULAR LOCATION.			
RP	MEDLINE=99158179; PubMed=10050892;			
RX	Dalmau J., Gultekin S.H., Voltz R., Hoard R., DesChamps T.,			
RA	Balmaceda C., Batchelor T., Gerstner E., Eichen J., Fremier J.,			
RA	Posner J.B., Rosenfeld M.R.;			
RT	"Ma1, a novel neuron- and testis-specific protein, is recognized by the serum of patients with paraneoplastic neurological disorders.";			
RT	Brain 122:27-39(1999).			
RL	Brain 122:27-39(1999).			
CC	-!- SUBCELLULAR LOCATION: Nuclear; nucleolar.			
CC	-!- TISSUE SPECIFICITY: Brain specific. In some patients suffering from cancers, it is also specifically expressed by the testicular tumor cells.			
CC	-!- MISCELLANEOUS: Antibodies against PMA2 are present in sera from patients suffering of paraneoplastic neurological disorders.			
CC	-!- SIMILARITY: Belongs to the PNMA family.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/			

042287 xenopus lae
P48725 mus musculus
P19246 mus musculus
P57660 buchnera ap
Q8k0d5 mus musculus
Q9yhy6 xenopus lae
Q06637 gallus gall
P15112 dictyosteli
P46504 caenorhabdi
Q9y006 mus musculus
Q08908 mus musculus
Q04637 homo sapien

ALIGNMENTS

34 98 1370 1 ITN1_XENLA
35 98 1320 1 PCT2_MOUSE
36 97 1087 1 NFH_MOUSE
37 96.5 6.6 611 1 HSCA_BUCAI
38 96.5 6.6 751 1 EFG1_MOUSE
39 96.5 6.6 1364 1 CND1_XENLA
40 96 6.6 657 1 BPS1_CHICK
41 96 6.6 830 1 EF2_DICDI
42 96 6.6 887 1 YLX8_CABEL
43 96 6.6 2114 1 MY9B_MOUSE
44 95.5 6.5 722 1 P85B_MOUSE
45 95.5 6.5 1395 1 IF4G_HUMAN

```
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF037365; AAD02098.1; -
DR EMBL; AF083114; AAF05625.1; -
DR EMBL; AF083115; AAF05626.1; -
DR EMBL; AF286487; AAG28165.1; -
DR EMBL; AB020690; BAA74906.1; -
DR GenBank; HGNC:9159; PMA2.
KW Antigen; Tumor antigen; Nuclear protein.
FT DOMAIN 333 338 POLY-GLU.
FT CONFLICT 128 129 GV -> AL (IN REF. 1; AAD02098).
FT CONFLICT 141 141 LV -> I (IN REF. 1; AAD02098).
FT CONFLICT 257 257 T -> P (IN REF. 1; AAF05625).
FT CONFLICT 278 278 R -> K (IN REF. 1; AAF05626).
SQ SEQUENCE 364 AA; 41509 MW; 6E417AD96B3F0E93 CRC64;

Query Match 99.8%; Score 1459; DB 1; Length 364;
Best Local Similarity 99.6%; Pred. No. 1.1e-96;
Matches 282; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VQGGGVKVIPTKPNQDTEFLERLNLFLEKGGQTVSGMFRALGQGVSPATVPCISPEL 60
Db 82 VQGGGVKVIPTKPNQDTEFLERLNLFLEKGGQTVSGMFRALGQGVSPATVPCISPEL 141
QY 61 LAHLIGQAWAHAPQPLPMRYKLVFSGSAPAPPEESFEVWLEQATEIVKEWVPVTEAE 120
Db 142 LAHLIGQAWAHAPQPLPMRYKLVFSGSAPAPPEESFEVWLEQATEIVKEWVPVTEAE 201
QY 121 KKRWLAEISLRGPAIDLMHIVQADNPSISVEECLEAFKQVFGSLESRRTAQVRLKTYQEE 180
Db 202 KKRWLAEISLRGPAIDLMHIVQADNPSISVEECLEAFKQVFGSLESRRTAQVRLKTYQEE 261
QY 181 GKVSAYVLRLETLRKAVKRAIPRIADQVRLQVMAQATLNQMLWCLRELKQGGPP 240
Db 262 GKVSAYVLRLETLRKAVKRAIPRIADQVRLQVMAQATLNQMLWCLRELKQGGPP 321
QY 241 PSFLELMKVIREEESEAFSENEISIEPEERDGYGRWNHEGDD 283
Db 322 PSFLELMKVIREEESEAFSENEISIEPEERDGYGRWNHEGDD 364

RESULT 2
ID PMA2_MACFA STANDARD; PRT; 364 AA.
AC Q9GMC3;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DE Paraneoplastic antigen Ma2 homolog.
GN PMA2.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OC NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
RA Suzuki Y., Sugano S., Hashimoto K.;
RT "Isolation of full-length cDNA clones from macaque brain cDNA
RL libraries."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Nuclear; nucleolar (By similarity).
CC -!- SIMILARITY: Belongs to the PMA family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
```


RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shemmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mollahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalilus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [4]
RP SEQUENCE OF 162-353 FROM N.A.
RC TISSUE=Testis;
RL Poustka A., Wellenreuther R., Mewes H.-W., Weil B., Wiemann S.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Nuclear; nucleolar. In tumor cells, it is
CC cytoplasmic.
CC -!- TISSUE SPECIFICITY: Testis and brain specific. In some patients
CC suffering from cancers, it is also specifically expressed by the
CC paraneoplastic tumor cells.
CC -!- MISCELLANEOUS: Antibodies against PNMAL are present in sera from
CC patients suffering of paraneoplastic neurological disorders.
CC -!- SIMILARITY: Belongs to the PNMA family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF037364; AAD13810.3; -
CC EMBL; AF320308; AAN05100.1; -
CC EMBL; BC039577; AAH39577.1; -
CC EMBL; AL834327; CAD38995.1; -
CC Genew; HGNC:9158; PNMA1.
CC MIM; 604010; -
CC GO; GO:0005737; C:cytoplasm; TAS.
CC GO; GO:0005730; C:nucleolus; TAS.
CC GO; GO:0007417; P:central nervous system development; TAS.
CC GO; GO:0007283; P:spermatogenesis; TAS.
CC Antigen; Tumor antigen; Nuclear protein.
CC DOMAIN 335 341 POLY-GLU.
CC SEQUENCE 353 AA; 39761 MW; EB7F5B6AEDA25961 CRC64;
CC
CC Query Match 43.08; Score 628; DB 1; Length 353;
CC Best Local Similarity 47.9%; Pred. No. 1.2e-37;
CC Matches 128; Conservative 52; Mismatches 75; Indels 12; Gaps 4;
CC
CC QY 3 GKGGVKVIKFTPNQDTFLERLNLFLKEGQTVSGMFRALGQGVSPATVPCISPPELLA 62
CC DB 83 GKGGVKVLKPKPTSDAEFLERLHLFLAREGTVQDVARVLGFGNPFTPTP-----GPEMPA 138
CC
CC QY 63 HLLGQMAHAPOPLL-PMRYKRLRVFGSAVPAPEESFEVWLEQATEIVKEHPVTEAK 121
CC DB 139 EMLNYILDNIQPLVESIWIYKRLTLFSGRDIPGSGEETFDPLWLEHTNEVLEEQVSVDEK 198
CC
CC QY 122 KRWLAESIRGALDMLHIVQADNPSISVBECLFAKQVFGSLSESRRTAQVRYLKYTOEEG 181
CC DB 199 RRLMESLRGPAADVIRILKSNPAITTAECCKALEQVFGSVESDQAQKFLNTYQNP 258
CC
CC QY 182 EKVSAYVLRLETLRLKAVEKAIPIRIADQVRLEQVMAGA-----TLNQMLWCRLELKDQ 237

Db 259 EKLSAYVIRLEPLLOKVVKEGAIDKDNVNQARLEQVIAHNSGAIRQLWL---TGAGE 315
QY 238 GPPPSFLMKVIREEEEBASPENES 264
Db 316 GPAPNLFOLLVQIREEEAKEEEAEAE 342

RESULT 6
PMAL MOUSE
ID PMAL MOUSE STANDARD; PRT; 353 AA.
DT 08CIC8; Q9CYP2;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Paraneoplastic antigen Mal homolog.
GN PNMA1 OR MAL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Brain, and Embryo;
RX MEDLINE=22354683; PubMed=12466851;
RA Okazaki Y., Furuno M., Sakukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Perlea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlstedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wyszynski B., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
CC -!- SUBCELLULAR LOCATION: Nuclear; nucleolar (By similarity).
CC -!- SIMILARITY: Belongs to the PNMA family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AK017476; BAB30762.1; -
CC EMBL; AK028331; BAC25885.1; -
CC MGD; MGI:2180564; Pmal.
CC Nuclein protein.
CC DOMAIN 336 341 POLY-GLU.
CC CONFLICT 136 136 T -> M (IN REF. 1; BAB30762).
CC SEQUENCE 353 AA; 39688 MW; 18CEDC3AC4E70939 CRC64;
CC

Query Match 42.2%; Score 617; DB 1; Length 353;
 Best Local Similarity 49.1%; Pred. No. 7.3e-37;
 Matches 130; Conservative 47; Mismatches 72; Indels 16; Gaps 5;

QY 3 GKGWGVKVFPTNODTEFLRLNLFLEKEGTVSGMFRALQGEVSPATVPCISPPELLA 62
 DB 83 GKGGLMKVVFPTSDAVFLERLHLFLAREGTVQDVARVLGFQNPAPAP-----GPETPA 138
 QY 63 HLLGQAMAHAPQPLP-PMYRKLRFVSGSAVPAPPEESFEVWLEQATEIVKEWPTAEK 121
 DB 139 EMLNYILDNVLOPLVESIWKYKLTUFGKNDIPGGEETFDWSLHSENIIEBWQSDIEK 198
 QY 122 KRWLAESLRGPDALDHIVQADNPISVIEECLEAFKQVFGSLERRTAQVRVLYKTYQEG 181
 DB 199 RRLMESLRGPAADVIRILKTNAAITTAECLEKALEQVFGSVESRDAQVRFNLTYQNGP 258
 QY 182 EKVSAYVLRLETLRLKAVEKRAIPRIADQVRLQVWAGAE-----TLNQLMWCRLRLK 237
 DB 259 EKLSSYVIRLEPLQKVVVDKGVIDKDNVNOARLEQVAGNHSGLRRLQMLAGAE---E 315
 QY 238 GPPPSFLELMKVIREE-----EEEA 258
 DB 316 GPAENLFQLLVIREEAKKEEEA 340

RESULT 7
 MOPI_HUMAN
 ID MOPI_HUMAN STANDARD; PRT; 351 AA.
 AC Q96BY2; Q9H833; Q9HAS1;
 DT 15-MAR-2004 (Rel. 43, Created)
 DT 15-MAR-2004 (Rel. 43, Last sequence update)
 DE Modulator of apoptosis 1 (MAP-1) (Paraneoplastic antigen MA4).
 GN MAP1 OR PNMA4
 GN Homo sapiens (Human).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., CHARACTERIZATION, AND MUTAGENESIS OF LEU-120;
 RC 120-LEU-ARG-127 AND 125-GLY-GLU-127.
 RP TISSUE=Cerebellum;
 RX MEDLINE=21264738; PubMed=11060313;
 RA Tan K.O., Tan K.M.L., Chan S.-L., Yee K.S.Y., Bevoit M., Ang K.C.,
 RA Yu V.C.;
 RT "MAP-1, a novel proapoptotic protein containing a BH3-like motif that
 RT associates with Bax through its Bcl-2 homology domains.";
 RL J. Biol. Chem. 276:2802-2807(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Retinoblastoma;
 RA Isoqai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
 RA Nishikawa T., Nagai K., Sato H., Sugano S., Shiratori A., Sudo H.,
 RA Wgatsuma M., Hosoari T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
 RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
 RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
 RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahara K., Masuho Y.,
 RA Ninomiya K., Iwayanagi T.;
 RT "NEDO human cDNA sequencing project.";
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Skin;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heise F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udén T.B., Toshikiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettman M., Madao A.C., Rodrigues S., Sanchez A.,
 RA Whiting M., Madao A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalusz D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -!- FUNCTION: Mediates caspase-dependent apoptosis.
 CC -!- SUBUNIT: Homodimer. Binds BAX, BCL2 and BCLX.
 CC -!- TISSUE SPECIFICITY: Ubiquitously expressed with higher levels in
 CC heart and brain.
 CC -!- DOMAIN: The BH3-like domain is required for association with BAX
 CC and for mediating apoptosis. The three BH domains (BH1, BH2, and
 CC BH3) of BAX are all required for mediating protein-protein
 CC interaction.
 CC -!- SIMILARITY: Belongs to the BAX family.
 CC -!- CAUTION: Ref.2 sequence differs from that shown due to a
 CC frameshift in position 102.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF305550; AAG31786.1; --
 DR EMBL; AK024029; BAB14788.1; ALT_SEQ.
 DR EMBL; BC015044; AAB15044.1; --
 DR Gene; HGNC:16658; MOAP1.
 KW Apoptosis.
 FT SITE 120 127 BH3-LIKE.
 FT MUTAGEN 120 127 MISSING: ABROGATED INTERACTION WITH BAX,
 FT MUTAGEN 120 127 RESULTING IN A NONAPOPTOTIC PROTEIN.
 FT MUTAGEN 120 120 L->E: WEAKENED INTERACTION WITH BAX,
 FT MUTAGEN 125 127 RESULTING IN A NONAPOPTOTIC PROTEIN.
 FT MUTAGEN 125 127 GHE->VLA: ABROGATED INTERACTION WITH BAX,
 FT CONFLICT 244 244 T -> A (IN REF. 2).
 FT CONFLICT 258 258 Y -> H (IN REF. 2).
 FT CONFLICT 259 259 Q -> H (IN REF. 1).
 SQ SEQUENCE 351 AA; 39512 MW; 5310142AC02B563C CRC64;

Query Match 40.8%; Score 596.5; DB 1; Length 351;
 Best Local Similarity 48.3%; Pred. No. 2.1e-35;
 Matches 128; Conservative 49; Mismatches 75; Indels 13; Gaps 5;

QY 1 VQKGGVWVIFKTPNODTEFLRLNLFLEKEGTVSGMFRALQGEVSPATVPCISP 60
 DB 82 IPKGGGVRVIFKPPDPDNTFLSRNLFELAGEGTVGLSPALCHENGSLDPEQMP 141
 QY 61 LAHLLGQAMAHAPQPLP-MRYRKLRFVSGSAVPAPPEESFEVWLEQATEIVKEWPT 119
 DB 142 WAPMLAQAAL-EALQPALQCLKYLRFVSGRESPEGESEFGFRWFFHTTQMIKA 200
 QY 120 EKKWLAESLRGPDALDHIVQADNPISVIEECLEAFKQVFGSLERRTAQVRVLYKTY 179
 DB 201 EKRRLESRGPDALDVIRILKTNAAITTAECLEKALEQVFGSVESRDAQVRFNLTY 260
 QY 180 EGEKVSAYVLRLETLRLKAVEKRAIPRIADQVRLQVWAGAE-----TLNQLMWCRLRLK 236
 DB 261 DEEKL SAYVLRLEPLQKLVQRAIEDAVNQARLDQVIAGAVHKTIRREL-----NLPE 315
 QY 237 QGPPPSFLELMKVIREE-----EEEA 258
 DB 316 DGPAPGFLQLLVILKDYEAEEEA 340

RESULT 8

```

MOPI1_MACPA
ID MOPI1_MACPA STANDARD; PRT; 351 AA.
AC Q95KI4;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Modulator of apoptosis 1 (MAP-1).
GN MOAP1.
OS Macaca fascicularis (Crab eating macaque). (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Temporal cortex;
RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
RA Suzuki Y., Sugano S., Hashimoto K.;
RT "Isolation of full-length cDNA clones from macaque brain cDNA
RT libraries";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Mediates caspase-dependent apoptosis (By similarity).
CC -!- SUBUNIT: Homodimer. Binds BAX, BCL2 and BCLX (By similarity).
CC -!- DOMAIN: The BH3-like domain is required for association with BAX
CC and for mediating apoptosis. The three BH domains (BH1, BH2, and
CC BH3) of BAX are all required for mediating protein-protein
CC interaction (By similarity).
CC -!- SIMILARITY: Belongs to the PNMA family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to licenses@isb-sib.ch).
CC -----
DR EMBL; AB060854; BAB46873.1; -
KW Apoptosis.
FT SITE 120 127 BH3-LIKE.
FT DOMAIN 336 339 POLY-GLU.
SQ SEQUENCE 351 AA; 39623 MW; C7530E4496A6FFB3 CRC64;

Query Match 40.6%; Score 594; DB 1; Length 351;
Best Local Similarity 47.1%; Pred. No. 3.1e-35;
Matches 128; Conservative 51; Mismatches 81; Indels 12; Gaps 5;

QY 1 VQKGGVWVKVIFKTPNQDTEFLERLNFLEKEQTVSGMFRALQGVSPATVPCISPEL 60
DB :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
82 IFPGKGIWRVIFKPPDSNTFLSLNEFLAGEQMTVGLTRALAHENGSLDLFQGMIPEM 141
QY 61 LAHLQGMAMHAPQPLP-MRYTKLVFSGSAVPAPPEESFEVWLQATEIVKWEVPTFA 119
DB :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
142 WAPMLAQAL-EALQPALQCLIKYKLLRVFSGREPPEFGEFGRWMTHTQMKAKAWQVDPV 200
QY 120 EKRWLAESLGRGALDMLHIVQADNPISIVEECLAFKQVFGSLERRTAQVRVLYKTYOE 179
DB :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
201 EKRRLLSGLRGALDVRIVRLKNNPLITVDECLQALEVFGVTDNPRELQVLYITYQK 260
QY 180 EGEKVSAVYLRLTTLRKAYEKRAIPRIADQVRLEQVWAGA---TLNQLMWLCRLKLD 236
DB :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
261 DEEKL SAYVLRLPFLQKLVRGAIERDAVNQARLDQVIAGVHKVIRREL-----NLPE 315
QY 237 QGPPPSFLELMKVIRREE--EEASPENESIE 266
DB :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
316 DGPAPGFLQLLVLIKQYEAABEALLQEVLE 347

RESULT 9
MOPI_MOUSE
ID MOPI_MOUSE STANDARD; PRT; 352 AA.
AC Q9ERH6;

```

```

DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Modulator of apoptosis 1 (MAP-1).
GN MOAP1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=21264738; PubMed=11060313;
RA Tan K.O., Tan K.M.L., Chan S.-L., Yee K.S.Y., Bevort M., Ang K.C.,
RA Yu V.C.;
RT "MAP-1, a novel proapoptotic protein containing a BH3-like motif that
RT associates with Bax through its Bcl-2 homology domains.";
RL J. Biol. Chem. 276:2802-2807(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX STRAIN=C57BL/6J; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez L., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6NCr; TISSUE=Hematopoietic, and Kidney;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uddin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gnarinate P.H.,
RA Richards S., Wreny K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: Mediates caspase-dependent apoptosis.
CC -!- SUBUNIT: Homodimer. Binds BAX, BCL2 and BCLX (By similarity).
CC -!- DOMAIN: The BH3-like domain is required for association with BAX
CC and for mediating apoptosis. The three BH domains (BH1, BH2, and
CC BH3) of BAX are all required for mediating protein-protein
CC interaction (By similarity).
CC -!- SIMILARITY: Belongs to the PNMA family.

```

 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

 EMBL: U83919; AAB41543.1; -
 InterPro: IPR006821; Filament_head.
 InterPro: IPR001664; IF.
 Pfam: PF00038; filament; 1.
 Pfam: PF04732; filament head; 1.
 PROSITE: PS00226; IF; 1.
 Intermediate filament; Coiled coil; Neurone.
 INIT MET 0 BY SIMILARITY.
 FT DOMAIN 1 92 HEAD (BY SIMILARITY).
 FT DOMAIN 93 396 ROD (BY SIMILARITY).
 FT DOMAIN 397 554 TAIL (BY SIMILARITY).
 FT DOMAIN 93 124 COIL 1A.
 FT DOMAIN 125 137 LINKER 1.
 FT DOMAIN 138 233 COIL 1B.
 FT DOMAIN 234 252 LINKER 12.
 FT DOMAIN 253 271 COIL 2A.
 FT DOMAIN 272 280 LINKER 2.
 FT DOMAIN 281 396 COIL 2B.
 FT DOMAIN 397 443 TAIL, SUBDOMAIN A.
 FT DOMAIN 444 554 TAIL, SUBDOMAIN B (ACIDIC).
 FT CONFLICT 494 500 MISSING (IN REF. 2).
 FT CONFLICT 509 509 A -> AEA (IN REF. 2).
 SQ SEQUENCE 554 AA; 62514 MW; D772B81CA2C31C1A CRC64;

Query Match 7.7%; Score 112; DB 1; Length 554;
 Best Local Similarity 24.2%; Pred. No. 0.97;
 Matches 55; Conservative 30; Mismatches 82; Indels 60; Gaps 10;

Qy 100 FEVWLEQA---TEIVKEMPVTEAEKKWL-----AESLRG--PALD--LMHIVQADNPS 146
 Db 284 FTVLTESAANKTDVAAKQDVESRRLLKAKTLEIEACRGWNALEKQLOLEDKQAD 343
 Qy 147 ISVEECLEAFKQVFGSLERRTAQVRYLKYOE-----EGEK----- 183
 Db 344 ISAMQ---DTINKLENELRTTKSEMARVLYKEYQDLINVKQALDIEIAAVRKLEGEETRLS 401
 Qy 184 -----VSAYVLRLETLRKA-----VKRAIPRIADQVRLQVWAGATLNQM 226
 Db 402 FTVSGSLTTGTQSQVFGSAYGGLQTSVYLSARSFPSTYTSVHVEQIEVEETIEA- 460
 Qy 227 LWCRLRELKQGPFPFLELMKLVIREEEERASFENESIEPEERDQ 273
 Db 461 --AKAEAKOE--PPEGEAEKEKEKEAEAEAEAEAEAEAEAEAE 503

RESULT 12
 NFL XENLA
 ID NFL XENLA STANDARD; PRT; 544 AA.
 AC P35616;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, last sequence update)
 DT 10-OCT-2003 (Rel. 42, last annotation update)
 DE Neurofilament triplet L protein (Neurofilament light polypeptide) (NF-L).
 DE L.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RT TISSUE=Brain;
 RX MEDLINE=92356194; PubMed=1494944;
 RA Charnas L.R., Szaro B.G., Gainer H.;

RT Identification and developmental expression of a novel low molecular weight neuronal intermediate filament protein expressed in Xenopus laevis.";
 RL J. Neurosci. 12:3010-3024(1992).
 CC -!- FUNCTION: Neurofilaments usually contain three intermediate filament proteins: L, M, and H which are involved in the maintenance of neuronal caliber.
 CC -!- MISCELLANEOUS: NF-L IS THE MOST ABUNDANT OF THE THREE NEUROFILAMENT PROTEINS AND, AS THE OTHER NONEPITHELIAL INTERMEDIATE FILAMENT PROTEINS, IT CAN FORM HOMOPOLYMERIC 10-NM FILAMENTS.
 CC -!- SIMILARITY: Belongs to the intermediate filament family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC

EMBL: M86654; AAB3018.1; -
 InterPro: IPR006821; Filament_head.
 InterPro: IPR001664; IF.
 Pfam: PF00038; filament; 1.
 Pfam: PF04732; filament_head; 1.
 PROSITE: PS00226; IF; 1.
 Intermediate filament; Coiled coil; Neurone.
 FT DOMAIN 1 87 HEAD.
 FT DOMAIN 88 390 ROD.
 FT DOMAIN 391 544 TAIL.
 FT DOMAIN 88 119 COIL 1A.
 FT DOMAIN 120 132 LINKER 1.
 FT DOMAIN 133 228 COIL 1B.
 FT DOMAIN 229 246 LINKER 12.
 FT DOMAIN 247 265 COIL 2A.
 FT DOMAIN 266 274 LINKER 2.
 FT DOMAIN 275 390 COIL 2B.
 FT DOMAIN 391 435 TAIL, SUBDOMAIN A.
 FT DOMAIN 436 544 TAIL, SUBDOMAIN B (ACIDIC).
 FT DOMAIN 441 538 GLU-RICH.
 FT DOMAIN 464 469 POLY-GLU.
 SQ SEQUENCE 544 AA; 61861 MW; 76D911B896E97201 CRC64;

Query Match 7.6%; Score 110.5; DB 1; Length 544;
 Best Local Similarity 22.3%; Pred. No. 1.2;
 Matches 59; Conservative 30; Mismatches 83; Indels 93; Gaps 10;

Qy 80 RYKRLRVFSGSAVPAPEESFEVWLE-----QATEIVKEMPVTEAEKKWLA-- 126
 Db 258 QYEKL-----AAKMQSAEWEFKSRFTVLTQSAARNTDAVRAAKDENSESRRLMSAK 309
 Qy 127 ----ESLRGPALDLMHIVQ--ADNPISVSEECLEAFKQVFGSLERRTAQVRYLKYOE- 179
 Db 310 GLEIEACRGVNEALQRQIQELEDKQSGEIAQMADINKLEELRNTKSEMARVLYKEYQDL 369
 Qy 180 -----EGEKV-----SAYVLRLETLRKA 198
 Db 370 LNVKQALDIEIAAYVRKLEGEETRLSFGVGAITSGYTSQAPVFGRSAYSLQSSYM--- 426
 Qy 199 VKRAIPRIADQVRLQVWAGATLNQMLWCLRELKQGPFPFLELMKLVIREEEEREA 258
 Db 427 -TSRAPFTYSSHVQEEQLDIEETIS---SRAEAKAEAP-----EEEEEA 471
 Qy 259 SFENESIEPEERDGYGRWNHGGD 283
 Db 472 EEEGEGEAEAE-----GEEGEE 490

RESULT 13
 ITN1 MOUSE
 ID ITN1 MOUSE STANDARD; PRT; 1714 AA.
 AC Q920R4; Q9R143;

28-FEB-2003 (Rel. 41, Created)
28-FEB-2003 (Rel. 41, Last sequence update)
10-OCT-2003 (Rel. 42, Last annotation update)
DE Intersectin 1 (EH and SH3 domains protein 1).
GN ITSN1 OR ITSN OR ESEL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RN SEQUENCE FROM N.A. (ISOFORM 1).
RP MEDLINE=99164083; PubMed=10064583;
RX Sengar A.S., Wang W., Bishay J., Cohen S., Egan S.E.;
RA "The EH and SH3 domain ESE proteins regulate endocytosis by linking to
RT dynam and Eps15.";
RL EMBO J. 18:1159-1171 (1999).
[2]
RN SEQUENCE OF 966-1714 FROM N.A. (ISOFORM 2), AND SEQUENCE OF 545-599
RP FROM N.A. (ISOFORMS 1 AND 2).
RC STRAIN=129/Ola; TISSUE=Spleen;
RA Skripkina I.Y., Tsyba L.O., Anoprienko O.V., Slavov D., Tassone F.,
RA Rynditch A.V., Gardiner K.;
RT "Mouse homologues of human chromosome 21 genes.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Adapter protein that may provide indirect link between
CC the endocytic membrane traffic and the actin assembly machinery.
CC May regulate the formation of clathrin-coated vesicles.
CC -!- SUBUNIT: Interacts with dynamin, SNAP-25 and SNAP-23. Clusters
CC several dynamin in a manner that is regulated by alternative
CC splicing. Also binds clathrin-associated proteins and other
CC components of the endocytic machinery, such as N-WASP, Eps15 and
CC Stoinin 2 (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic; membrane-associated protein.
CC Enriched in synaposomes (By similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=2;
CC Comment=Additional isoforms seem to exist;
CC Name=1; Synonyms=Esel;
CC IsoId=Q9Z0R4-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q9Z0R4-2; Sequence=VSP_004296;
CC -!- TISSUE SPECIFICITY: Widely expressed. Expressed at high levels in
CC brain, heart and skeletal muscle.
CC -!- DOMAIN: SH3-3, SH3-4 and SH3-5, but not SH3-1 and SH3-2 domains,
CC bind to dynamin (By similarity).
CC -!- DOMAIN: The KLERQ domain binds to SNAP-25 and SNAP-23 (By
CC similarity).
CC -!- MISCELLANEOUS: Overexpression results in the inhibition of the
CC transferrin uptake and the blockage of the clathrin-mediated
CC endocytosis.
CC -!- SIMILARITY: Contains 1 C2 domain.
CC -!- SIMILARITY: Contains 1 DBL-homology (DH) domain.
CC -!- SIMILARITY: Contains 2 EF-hand calcium-binding domains.
CC -!- SIMILARITY: Contains 2 EH domains.
CC -!- SIMILARITY: Contains 1 PH domain.
CC -!- SIMILARITY: Contains 5 SH3 domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF132481; AAD19749.1; -
CC EMBL; AF132478; AAD19746.1; -
CC EMBL; AF169621; AAD48848.1; -
CC EMBL; AF356517; AAK40228.1; -
CC HSSP; P29354; 1GFC
CC MGD; MGI:1338069; Itsn.
CC GO; GO:0030027; C:lamellipodium; IDA.
CC GO; GO:0005515; F:protein binding; IPI.

DR GO; GO:0007264; P:small GTPase mediated signal transduction; IDA.
DR InterPro; IPR000008; C2.
DR InterPro; IPR008973; C2_CaLB.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR000261; EFS15_homology.
DR InterPro; IPR001849; PH.
DR InterPro; IPR000219; RhoGEF.
DR InterPro; IPR001452; SH3.
DR Pfam; PF00168; C2_1.
DR Pfam; PF00036; ehfand; 3.
DR Pfam; PF00169; PH; 1.
DR Pfam; PF00621; RhoGEF; 1.
DR Pfam; PF00018; SH3; 5.
DR PRINTS; PR00452; SH3DOMAIN.
DR ProDom; PD000066; SH3; 5.
DR SMART; SM00239; C2; 1.
DR SMART; SM00054; Efh; 2.
DR SMART; SM00027; EH; 2.
DR SMART; SM00233; PH; 1.
DR SMART; SM00325; RhoGEF; 1.
DR SMART; SM00326; SH3; 5.
DR PROSITE; PS50004; C2_DOMAIN_2; 1.
DR PROSITE; PS50010; DH_2; 1.
DR PROSITE; PS00018; EF_HAND; 2.
DR PROSITE; PS00031; EH; 2.
DR PROSITE; PS50003; PH_DOMAIN; 1.
DR PROSITE; PS50002; SH3; 5.
KW Endocytosis; SH3 domain; Repeat; Coiled coil; Calcium-binding;
KW Alternative splicing.
FT DOMAIN 21 109 EH 1.
FT CA_BIND 66 78 EF-HAND 1 (POTENTIAL).
FT DOMAIN 221 310 EH 2.
FT CA_BIND 267 279 EF-HAND 2 (POTENTIAL).
FT DOMAIN 326 702 LYS/LEU/GLU/ARG/GLN-RICH (KLERQ).
FT DOMAIN 352 662 COILED COIL (POTENTIAL).
FT DOMAIN 738 799 SH3 1.
FT DOMAIN 906 964 SH3 2.
FT DOMAIN 995 1053 SH3 3.
FT DOMAIN 1067 1131 SH3 4.
FT DOMAIN 1148 1207 SH3 5.
FT DOMAIN 1230 1416 DH.
FT DOMAIN 1455 1564 PH.
FT DOMAIN 1576 1672 C2 DOMAIN.
FT DOMAIN 321 324 POLY-SER.
FT VARSPLOC 1214 1714 Missing (in isoform 2).
FT SEQUENCE 1714 AA; 194284 MW; 4D7AF298397860A7 CRC64;
Query Match 7.6%; Score 110.5; DB 1; Length 1714;
Best Local Similarity 26.0%; Pred. No. 4.8;
Matches 60; Conservative 23; Mismatches 61; Indels 87; Gaps 9;
QY 60 LLAHLIGQAMAHAPQ-----LLPMRYKLRVFGSAVPA-----PEESFVWLE 105
DB 281 LAWHLIDVAMSGQPLDPVLPPEYIPPSFRVRGSGMSVSSSVQRIPEESSED--E 338
QY 106 QATEIVKVPVTEAEKKRWLAESLRGALDMLHIVQADNPSISVEECLEAFKQVFGSLES 165
DB 339 QQPE--KLPVTFEDKKR-----ENFERGSEVELEK 366
QY 166 RRTAQVRYLYKTQVEGEKVSAYVRLTLLRKAVEKRAIPRRITADQVRLQVWAGATLQ 225
DB 367 RQALLLEQQRKEQ-----RLAQLERAFQERK-----ERERQEQ----- 400
QY 226 MLWCRLRELKQGPFPSPFLELMKVIREEEEFEEFASFENESIEEPEEDDGVR 276
DB 401 -----EAKRQ-----LELEKQLEKQRELEKQREERKEEREAAR 438
RESULT 14
ITN1 RAT
ID ITN1 RAT STANDARD; PRT; 1217 AA.
AC Q9WVE9; Q9WVE1;

DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Intersectin 1 (EH domain and SH3 domain regulator of endocytosis 1).
 GN ITSN1 OR ITSN OR EHSH1.
 OS Rattus norvegicus (Rat).
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 RN NCBI_TaxID=10116;
 RX MEDLINE=99303609; PubMed=10373452;
 RA Okamoto M., Schoch S., Suedhof T.C.;
 RT "EHSH1/intersectin, a protein that contains EH and SH3 domains and
 RT binds to dynamin and SNAP-25. A protein connection between exocytosis
 RT and endocytosis?";
 RL J. Biol. Chem. 274:18446-18454 (1999).
 CC -!- FUNCTION: Adapter protein that may provide indirect link between
 CC the endocytic membrane traffic and the actin assembly machinery.
 CC May regulate the formation of clathrin-coated vesicles.
 CC -!- SUBUNIT: Interacts with dynamin, SNAP-25 and SNAP-23. Clusters
 CC several dynamin in a manner that is regulated by alternative
 CC splicing. Also binds clathrin-associated proteins and other
 CC components of the endocytic machinery, such as N-WASP, Eps15 and
 CC stonin 2 (By similarity).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic; membrane-associated protein.
 CC -!- Enriched in synaptosomes.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event-Alternative splicing; Named isoforms=2;
 CC Comment=Additional isoforms seem to exist;
 CC Name=1;
 CC IsoId=Q9WVE9-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=Q9WVE9-2; Sequence=VSP 004297;
 CC -!- TISSUE SPECIFICITY: Highly expressed in brain.
 CC -!- DOMAIN: SH3-3, SH3-4 and SH3-5, but not SH3-1 and SH3-2 domains,
 CC bind to dynamin.
 CC -!- DOMAIN: The KLRQ domain binds to SNAP-25 and SNAP-23.
 CC -!- SIMILARITY: Contains 2 EF-hand calcium-binding domains.
 CC -!- SIMILARITY: Contains 2 EH domains.
 CC -!- SIMILARITY: Contains 5 SH3 domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF127798; AAD30271.1; -;
 DR EMBL; AF132672; AAD31026.1; -;
 DR HSP; P29354; IGFC.
 DR InterPro; IPR002048; EF-hand.
 DR InterPro; IPR000261; EPS15_homology.
 DR InterPro; IPR001452; SH3.
 DR Pfam; PF00036; efhand; 3.
 DR Pfam; PF00018; SH3; 5.
 DR PRINTS; PD00452; SH3DOMAIN.
 DR ProDom; PD000066; SH3; 5.
 DR SMART; SM00054; EFh; 2.
 DR SMART; SM00027; EH; 2.
 DR SMART; SM00326; SH3; 5.
 DR PROSITE; PS00018; EF_HAND; 2.
 DR PROSITE; PS50031; EH; 2.
 DR PROSITE; PS50002; SH3; 5.
 DR Endocytosis; SH3 domain; Repeat; Coiled coil; Calcium-binding;
 KW Alternative splicing.
 FT DOMAIN 21 109 EH 1.
 FT CA_BIND 66 78 EF_HAND 1 (POTENTIAL).
 FT DOMAIN 221 310 EH 2.
 FT CA_BIND 267 279 EF_HAND 2 (POTENTIAL).

FT DOMAIN 326 702 LYS/LEU/GLU/ARG/GLN-RICH (KLRQ).
 FT DOMAIN 350 670 COILED COIL (POTENTIAL).
 FT DOMAIN 737 803 SH3 1.
 FT DOMAIN 910 968 SH3 2.
 FT DOMAIN 999 1057 SH3 3.
 FT DOMAIN 1071 1135 SH3 4.
 FT DOMAIN 1152 1211 SH3 5.
 FT DOMAIN 321 324 POLY-SER.
 FT VARSPLIC 1003 1073 Missing (in isoform 2).
 FT SEQUENCE 1217 AA; 137154 MW; 6C13238AE5A5B34B CRC64;
 SQ SEQUENCE 1217 AA; 137154 MW; 6C13238AE5A5B34B CRC64;
 Query Match 7.4%; Score 108.5; DB 1; Length 1217;
 Best Local Similarity 24.7%; Pred. No. 4.4;
 Matches 57; Conservative 25; Mismatches 62; Indels 87; Gaps 8;
 QY 60 LLAHLGLQAMAHAPQ-----LLPMRYRKLKRVFSGSAVPA-----PEESFEVWLE 105
 Db 281 LAMHLIDVAMSGQPLPPVLPPEYIPSPFRVRSGSGMSVSSSSADQRLPEPS-----SE 336
 QY 106 QATEIVKWPVTAEAKKRWLAESLURGPALDLMHIVQADNPISVSECELEAFKQVFGSLES 165
 Db 337 DEQVKEKKLPVTFEDKKR-----ENFERGNLELEK 366
 QY 166 RRTAQVRLKTYOEGEKVSAYVLETLRLKXAVEKRAIPRIADQVLEQVMAGATLNQ 225
 Db 367 RQALLLEQQRKEQ-----RLAQLEAERK-----ERERQEQ 400
 QY 226 MLWCLRELKDGPPPSFLMKVIREEEEAAGFENESIEPERDGYGR 276
 Db 401 -----ERKEQ-----LELEKLEKQRELEKQREERKEIERREAAKR 438
 RESULT 15
 NAB2 MOUSE
 ID NAB2 MOUSE STANDARD; PRT; 525 AA.
 AC Q61127;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE NGFI-A binding protein 2 (EGR-1 binding protein 2).
 GN NAB2.
 OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 [1]
 RN SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 RP TISSUE=Brain;
 RX MEDLINE=96251303; PubMed=8668170;
 RA Svaren J., Severson B.R., Apel E.D., Zimonjic D.B., Popescu N.C.,
 RA Milbrandt J.;
 RT "NAB2, a corepressor of NGFI-A (Egr-1) and Krox20, is induced by
 RT proliferative and differentiative stimuli";
 RL Mol. Cell. Biol. 16:3545-3553 (1996).
 RN [2]
 RP PARTIAL SEQUENCE FROM N.A. (ISOFORM 2).
 RX MEDLINE=97271553; PubMed=9126479;
 RA Svaren J., Apel E.D., Simburger K.S., Jenkins N.A., Gilbert D.J.,
 RA Copeland N.A., Milbrandt J.;
 RT "The Nab2 and Stat6 genes share a common transcription termination
 RT region";
 RL Genomics 41:33-39 (1997).
 CC -!- FUNCTION: Acts as a transcriptional repressor for zinc finger
 CC transcription factors EGR1 and EGR2. Isoform 2 lacks repression
 CC ability.
 CC -!- SUBUNIT: Homomultimeris may associate with EGR1 bound to DNA (By
 CC similarity).
 CC -!- SUBCELLULAR LOCATION: NUCLEAR. ISOFORM 2 IS NOT LOCALIZED TO THE
 CC NUCLEUS.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event-Alternative splicing; Named isoforms=2;
 CC Name=1;

```

CC      Isoid=Q61127-1; Sequence=Displayed;
CC      Name=2;
CC      Isoid=Q61127-2; Sequence=VSP_003388, VSP_003389;
CC      TISSUE SPECIFICITY: Highly expressed in brain and thymus, and at
CC      lower levels in spleen, kidney, heart and testis. Isoform 1 is
CC      predominantly expressed in testis, whereas isoform 3 is more
CC      abundant in thymus.
CC      !- INDUCTION: By serum stimulation.
CC      !- DOMAIN: The NAB conserved domain 1 (NCD1) interacts with EGR1
CC      inhibitory domain and mediates multimerization.
CC      !- DOMAIN: The NAB conserved domain 2 (NCD2) is necessary for
CC      transcriptional repression.
CC      !- SIMILARITY: BELONGS TO THE NAB FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL; U47543; AAC52650.1; -
CC      MGD; MGI:107563; Nab2.
CC      InterPro; IPR006989; Nab centr.
CC      InterPro; IPR006988; Nab_N.
CC      Pfam; PF04904; NCD1; 1.
CC      Pfam; PF04905; NCD2; 1.
CC      Transcription regulation; Repressor; Alternative splicing.
CC      FT DOMAIN 35 113 NCD1.
CC      FT DOMAIN 267 356 NCD2.
CC      FT DOMAIN 353 384
CC      FT VARSPLIC 320 322 LTI -> ASL (in isoform 2).
CC      FT VARSPLIC 323 525 /FTId=VSP_003388.
CC      FT VARSPLIC 323 525 Missing (in isoform 2).
CC      FT /FTId=VSP_003389.
CC      SQ SEQUENCE 525 AA; 56653 MW; FED428E94A8BD804 CRC64;
CC      -----
Query Match 7.3%; Score 107; DB 1; Length 525;
Best Local Similarity 22.6%; Pred. No. 2.1;
Matches 47; Conservative 39; Mismatches 84; Indels 38; Gaps 6;
QY 86 VFSGSAVPAP---EESFEVWLQATEIVKEWPTAEKKRWLAESLRGPALDLMHVOA 142
Db ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
225 VAAGGAGGCDRLPEPWVWVSVESVERIFPSFPRGTGETASLLKLNKLARSVGHIFEM 284
QY 143 DNPISIVVECLEAFKQVFGSLESRRTAQVRYLKYTOEKEKVS-----AYVLR 191
Db ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
285 DDHDAQKEEBIRKYSVIYGRLDSCR-----REGKQLSLHELITINEAAAFQCMRD 333
QY 192 ETLLKAVEKRAIPRIADQVRLQVWAGNTLN--QMLWCRLELKD-----QGP 239
Db ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
334 NTLRLRVELSLRSQVARESTYLSLKGSLRSLSEELGGPPLKKLQKEVGEQSHNEIQP 393
QY 240 PPSFLELMKVIREE-EEESASFENESIE 266
Db ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
394 PPGPESYAPPYRPSLEEDSASLSGESLD 421

```

Search completed: September 21, 2004, 13:32:39
 Job time : 9.71499 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 13:20:22 ; Search time 65.5774 Seconds
(without alignments)
1361.621 Million cell updates/sec

Title: US-10-037-860-11

Perfect score: 1462

Sequence: 1 VQGGGVWVKVIFKTPNQDTE.....SIEPEERDGYGRWNHGGD 283

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1462	100.0	283	4 Q9UL42	Q9ul42 homo sapien
2	1459	99.8	364	4 Q94959	Q94959 homo sapien
3	1428	97.7	364	6 Q9GMJ3	Q9gmj3 macaca fasc
4	1193.5	81.6	365	11 Q8BHKO	Q8bhk0 mus musculu
5	755	51.6	149	4 Q9UL43	Q9ul43 homo sapien
6	634	43.4	353	11 Q8VH24	Q8vh24 rattus norv
7	628	43.0	353	4 Q95144	Q95144 homo sapien
8	628	43.0	353	4 Q8NG07	Q8ng07 homo sapien
9	622	42.5	466	11 Q8JZW8	Q8jzw8 mus musculu
10	620	42.4	353	11 Q9CYP2	Q9cyp2 mus musculu
11	618.5	42.3	455	4 Q9H0A4	Q9h0a4 homo sapien
12	618.5	42.3	463	4 Q9UL41	Q9ul41 homo sapien
13	617	42.2	353	11 Q8CIC8	Q8cic8 mus musculu
14	597	40.8	448	4 Q8NET3	Q8net3 homo sapien
15	597	40.8	452	4 Q96PV4	Q96pv4 homo sapien
16	596.5	40.8	351	4 Q96BY2	Q96by2 homo sapien

```

17 594 40.6 351 6 Q95KI4
18 591.5 40.5 351 4 Q9HAS1
19 576 39.4 194 4 Q95145
20 571.5 39.1 352 11 Q9ERH6
21 480.5 32.9 237 4 Q9H833
22 475.5 32.5 399 4 Q96A40
23 453.5 31.0 192 4 Q8ND90
24 403.5 27.6 402 11 Q9CZA5
25 394 26.9 403 4 Q8TE36
26 391.5 26.8 393 11 Q9DB17
27 391.5 26.8 393 11 Q8VD24
28 380.5 26.0 402 4 Q8N1C1
29 147 10.1 430 11 Q80VM8
30 137.5 9.4 246 11 Q8C533
31 135 9.2 378 4 Q8N3H4
32 135 9.2 435 4 Q8GV59
33 117 8.0 538 4 Q9H0W5
34 117 8.0 538 4 Q8TB26
35 116.5 8.0 1220 13 Q8JFT5
36 116.5 8.0 1721 13 Q8JFT4
37 111 7.6 1150 17 Q8U256
38 110 7.5 281 16 Q7VNB6
39 110 7.5 386 4 Q8TCR7
40 110 7.5 956 13 Q7ZW40
41 109.5 7.5 1202 13 Q7ZYU9
42 107 7.3 525 11 Q80VR9
43 107 7.3 543 4 Q8IU72
44 107 7.3 627 4 Q7Z328
45 106 7.3 567 3 Q9UUS7

```

ALIGNMENTS

RESULT 1

```

Q9UL42 PRELIMINARY; PRT; 283 AA.
ID Q9UL42
AC Q9UL42;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE Paraneoplastic cancer-testis-brain antigen (Fragment).
GN MA4
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Gultekin S.H., Voltz R., Rosenfeld M.R., Gerstner E., Eichen J.,
RA Posner J.B., Dalmay J.;
RT "Identification of a novel cancer testis brain antigen using serum
RT antibodies from patients with testicular tumors and paraneoplastic
RT limbic encephalomyelitis.";
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF083115; AA05626.1; -.
DR InterPro; IPR005162; Retrotrans_gag.
DR Pfam; PF03732; Retrotrans_gag; 1.
FT NON TER 1
SQ SEQUENCE 283 AA; 32333 MW; E27BA6BCDDCD240A4 CRC64;

```

```

Query Match 100.0%; Score 1462; DB 4; Length 283;
Best Local Similarity 100.0%; Pred. No. 3.8e-107;
Matches 283; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 VQGGGVWVKVIFKTPNQDTEFLERLNLFLKEGQTVSGMRALQEGVSPATVPCISPEL 60
Db 1 VQGGGVWVKVIFKTPNQDTEFLERLNLFLKEGQTVSGMRALQEGVSPATVPCISPEL 60
Qy 61 LAHLGQMAHAQPQLLPMPYRKLRFVSGSAVPAPEEESFEVWLEQATEIVKEWPVTEAE 120
Db 61 LAHLGQMAHAQPQLLPMPYRKLRFVSGSAVPAPEEESFEVWLEQATEIVKEWPVTEAE 120

```

QY 121 KRWLAESLRGPAALDMLHIVQADNPSISVEECLEAFKQVFGSLERRTAQRVLYKTYQEE 180
DB 121 KRWLAESLRGPAALDMLHIVQADNPSISVEECLEAFKQVFGSLERRTAQRVLYKTYQEE 180
QY 181 GEKVSAYVLRLETLRKAVEKRAIPRRIADQVRLEQVMAGATLNQMLWCLRELKDKQGGP 240
DB 181 GEKVSAYVLRLETLRKAVEKRAIPRRIADQVRLEQVMAGATLNQMLWCLRELKDKQGGP 240
QY 241 PSFLELMKVIREEEBEASFENESIEBEPEERDGYGRWNHEGDD 283
DB 241 PSFLELMKVIREEEBEASFENESIEBEPEERDGYGRWNHEGDD 283

RESULT 2
Q94959 PRELIMINARY; PRT; 364 AA.
AC Q94959;
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)
DE Hypothetical protein KIAA0883 (Paraneoplastic associated brain-testis-cancer antigen).
DE cancer antigen).
GN KIAA0883.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=99156230; PubMed=10048485;
RA Nagase T., Ishikawa K., Suyama M., Kikuno R., Hirose M., Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XII. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro.";
RL DNA Res. 5:355-364(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA Eichen J.G., Dalmay J., Wade D., Rosenfeld M.R.;
RT "Characterization of a Brain-Testis-Cancer Antigen.";
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB020690; BAA74906.1; -
DR EMBL; AF286487; AAC28165.1; -
DR Genew; HGNC:9159; PNM2.
DR InterPro; IPR005162; Retrotrans_gag.
DR Pfam; PF03732; Retrotrans_gag; 1.
KW Hypothetical protein.
SQ SEQUENCE 364 AA; 41509 MW; 6E417AD96E3F0E93 CRC64;

Query Match 99.8%; Score 1459; DB 4; Length 364;
Best Local Similarity 99.8%; Pred. No. 9.1e-107; Mismatches 0; Indels 0; Gaps 0;
Matches 282; Conservative 1;

QY 1 VQKGKGVKVIKFTPNQDTEFLERLNLFLKEKGQTVSGMFRALQEGVSPATVPCISPEL 60
DB 82 VQKGKGVKVIKFTPNQDTEFLERLNLFLKEKGQTVSGMFRALQEGVSPATVPCISPEL 141
QY 61 LAHLQGAHAHQPLLPMRYKLRVFGSAVPAPEEESFEVWLEQATEIVKWPVTEAE 120
DB 142 LAHLQGAHAHQPLLPMRYKLRVFGSAVPAPEEESFEVWLEQATEIVKWPVTEAE 201
QY 121 KRWLAESLRGPAALDMLHIVQADNPSISVEECLEAFKQVFGSLERRTAQRVLYKTYQEE 180
DB 202 KRWLAESLRGPAALDMLHIVQADNPSISVEECLEAFKQVFGSLERRTAQRVLYKTYQEE 261
QY 181 GEKVSAYVLRLETLRKAVEKRAIPRRIADQVRLEQVMAGATLNQMLWCLRELKDKQGGP 240
DB 262 GEKVSAYVLRLETLRKAVEKRAIPRRIADQVRLEQVMAGATLNQMLWCLRELKDKQGGP 321
QY 241 PSFLELMKVIREEEBEASFENESIEBEPEERDGYGRWNHEGDD 283
DB 322 PSFLELMKVIREEEBEASFENESIEBEPEERDGYGRWNHEGDD 364

RESULT 3

Q9GMU3 PRELIMINARY; PRT; 364 AA.
AC Q9GMU3;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Hypothetical protein.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
OC Cercopitheidae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K., Suzuki Y., Sugano S., Hashimoto K.;
RT "Isolation of full-length cDNA clones from macaque brain cDNA libraries.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047632; BAB12156.1; -
DR InterPro; IPR005162; Retrotrans_gag.
DR Pfam; PF03732; Retrotrans_gag; 1.
KW Hypothetical protein.
SQ SEQUENCE 364 AA; 41350 MW; 0CF72210D7EC1524 CRC64;

Query Match 97.7%; Score 1428; DB 6; Length 364;
Best Local Similarity 97.2%; Pred. No. 2.5e-104; Mismatches 4; Indels 0; Gaps 0;
Matches 275; Conservative 4;

QY 1 VQKGKGVKVIKFTPNQDTEFLERLNLFLKEKGQTVSGMFRALQEGVSPATVPCISPEL 60
DB 82 VQKGKGVKVIKFTPNQDTEFLERLNLFLKEKGQTVSGMFRALQEGVSPATVPCISPEL 141
QY 61 LAHLQGAHAHQPLLPMRYKLRVFGSAVPAPEEESFEVWLEQATEIVKWPVTEAE 120
DB 142 LAHLQGAHAHQPLLPMRYKLRVFGSAVPAPEEESFEVWLEQATEIVKWPVTEAE 201
QY 121 KRWLAESLRGPAALDMLHIVQADNPSISVEECLEAFKQVFGSLERRTAQRVLYKTYQEE 180
DB 202 KRWLAESLRGPAALDMLHIVQADNPSISVEECLEAFKQVFGSLERRTAQRVLYKTYQEE 261
QY 181 GEKVSAYVLRLETLRKAVEKRAIPRRIADQVRLEQVMAGATLNQMLWCLRELKDKQGGP 240
DB 262 GEKVSAYVLRLETLRKAVEKRAIPRRIADQVRLEQVMAGATLNQMLWCLRELKDKQGGP 321
QY 241 PSFLELMKVIREEEBEASFENESIEBEPEERDGYGRWNHEGDD 283
DB 322 PSFLELMKVIREEEBEASFENESIEBEPEERDGYGRWNHEGDD 364

RESULT 4

Q8BHKO PRELIMINARY; PRT; 365 AA.
AC Q8BHKO;
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Similar to KIAA0883 protein.
GN PNM2 OR A830049P17RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cortex;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RA the RIKEN Genome Exploration Research Group Phase I & II Team;


```

RA Posner J.B., Rosenfeld M.R.;
RT "Wal, a novel neuron- and testis-specific protein, is recognized by
RL the serum of patients with paraneoplastic neurological disorders." ;
RL Brain 122:27-39(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Cerebellum;
RA Dalman J., Rosenfeld M.R., Voltz R., Hoard R.;
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
RL EMBL; AF037364; AAD13810.3 ; -
DR Genew; HGNC:9158; PNMA1.
DR GO; GO:0005737; C:cytoplasm; TAS.
DR GO; GO:0005730; C:nucleolus; TAS.
DR GO; GO:0007417; P:central nervous system development; TAS.
DR GO; GO:0007283; P:spermatogenesis; TAS.
SQ SEQUENCE 353 AA; 39800 MW; 3BB41691AE89AD3D CRC64;

Query Match 43.0%; Score 628; DB 4; Length 353;
Best Local Similarity 47.9%; Pred. No.2.4e-41;
Matches 128; Conservative 52; Mismatches 75; Indels 12; Gaps 4;

Qy 3 GKGWVKVIFKTPNQDTFFLRNLFLKEGQTVSGMFRALGQGVSPATVPCISPELLA 62
Db 83 GKGWVKVLFKPTSDAEFLRLHLFLAREGWTQDVVARVLFGQNPPTP---GPENPA 138
Qy 63 HLIQMAHAAPQLL-PMRYKRLRVFSGSAPVAPPEERSFEVWLQOATEIVKEWPVTEAK 121
Db 139 EMLNVILNVITQPLVESIWKYKELTLFSGRDIPGFGEEFTDFPWLHTNEVLEQVSDVEK 198
Qy 122 KRWLAESIRGPAALDMHIQVADNPISIVBERCLEAPKQVFGSLESRRTAOVRLYKTYQEEG 181
Db 199 RRRLMESLRGPAADVIRILKSNPNPAITTAECILKALEQVFGSVESSRDAQIKFLNTYQNP 258
Qy 182 EKVSAYVLRLETLTKFAVEKRAIPRIADQVRLQVMAGA---TLNQMLWCLRLKDXQ 237
Db 259 EKLSAYVLRLEPLLOKVVEKGAIKDQNVNQARLEQVIAGANHSIGAIRQLML---TGAGE 315
Qy 238 GPPSFLKMKVIREEEBEEASFENES 264
Db 316 GPAPNLPOLLVOIREEEAKEEEFEAEA 342

```

```

RESULT 8
Q8NG07
ID ID PRELIMINARY; PRT; 353 AA.
AC Q8NG07;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Paraneoplastic antigen.
GN PNMA1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Schueller M.M., Jenne D.E., Schutze dit Belkner N., Hohlfield R.,
RA Voltz R.;
RA "Mal (PNMA1).";
RT Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Prostate;
RA Strausberg R.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF320308; AAN05100.1; -.
DR EMBL; BC039577; AAH39577.1; -.
SQ SEQUENCE 353 AA; 39761 MW; EB7F5B6AEDA25961 CRC64;

Query Match 43.0%; Score 628; DB 4; Length 353;
Best Local Similarity 47.9%; Pred. No. 2.4e-41;
Matches 128; Conservative 52; Mismatches 75; Indels 12; Gaps 4;

```

Qy	3	KGKGVKVIKFTPNQDTFLERLNLFLEKGGQTVSGMFRALQGGVSGPATVPCISPELLA	62
Db	83	KGKGVKVLKFTPTSDAEFLERLHLFLAREGTVQDVARVLGQNPTPTD---GPEMPA	138
Qy	63	HLIGQAWAHAPQILL-PMRYRKLRVTVSGSNAVPAPEESFEVWLEQATEIVKEWPVTEAK	121
Db	139	EMLNYILDNIQPLVESIWKRLTLFSGRDIPGGEETFPWLEHTNEVLEEWQVSDVEK	198
Qy	122	KRWLASLRGALDLMHIVQADNPSTSVBECLAFKQVFGSLESRRTAQVRYLKYOEKG	181
Db	199	RRRLMESLRGPAADVIRILKSNPAITAECLKALEQVFGSVSSRDAQIKFLNTYQNP	258
Qy	182	EKVSAYVLRLETLIRKAVEKRAIPRIADQVRLEQVWAGA---TLNQMLWCRLRELKQ	237
Db	259	EKLSAYVIRLEPLQKVVEGAIDKDNVNQARLEQVIAAGHNSGAIIRQLWL---TGAGE	315
Qy	238	GPPPSFLELMKVIREEBEEAEAFENES	264
Db	316	GPAENLFQLLVQIREEEAEAEAEAE	342
RESULT 9			
Qy	QBUZW8	PRELIMINARY; PRT; 466 AA.	
AC	QBUZW8;		
DT	01-OCT-2002 (TrEMBLrel. 22, Created)		
DT	01-OCT-2002 (TrEMBLrel. 22, Last sequence update)		
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)		
DE	Similar to paraneoplastic antigen MA3.		
GN	PNMA3.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Retina;		
RA	Submitted (Aug-2002) to the EMBL/GenBank/DBJ databases.		
RL	EMBL; BC036726; AAH36726.1; -.		
DR	MGI; MGI:2180565; Pnma3.		
DR	GO; GO:0003824; F:catalytic activity; IEA.		
DR	GO; GO:0003676; F:nucleic acid binding; IEA.		
DR	GO; GO:0008152; P:metabolism; IEA.		
DR	InterPro; IPR001544; Aminotrans_IV.		
DR	InterPro; IPR001878; Znf CCHC.		
DR	Pfam; PF00098; zf-CCHC; 1.		
DR	PRINTS; PR00939; C2HCZNFINGER.		
DR	PROSITE; PS50158; ZF_CCHC; 1.		
DR	SEQUENCE 466 AA; 5404 MW; 113787E37B0B3AAD CRC64;		
Query Match			
Best Local Similarity 42.5%; Score 622; DB 11; Length 466;			
Matches 132; Conservative 55; Mismatches 72; Indels 8; Gaps 4			
Qy	1	VQKGGVWKVLEKFTPNQDTFLERLNLFLEKGGQTVSGMFRALQGGVSGPATVPCISPE	59
Db	82	TEGKGGPEVWVKKPHSDDEFLNHLHLEBERRTVSDMNRVLGHSHSPKTTISADF	141
Qy	60	-LLAHLGQAWAHAPQILL-PMRYRKLRVTVSGSNAVPAPEESFEVWLEQATEIVKEWPV	117
Db	142	WVAQTLGVM---QPLLEQMLYRELAVFSGNTTISFGLLAFLDSWLEHTTEMLQWQVP	197
Qy	118	EAEKKRWLAESLRGALDMLHIVQADNPSTSVBECLAFKQVFGSLESRRTAQVRYLKY	177
Db	198	EYEKRRRLMECLRGALQVNVNLRANNAITVKECLEALRQVFGSDNRKIAQLKPKAY	257
Qy	178	OEEGKVSAYVLRLETLIRKAVEKRAIPRIADQVRLEQVWAGATLNMQLWCRLRELKQ	237
Db	258	QBPFGKVSFVVRLETLIQKALEKNAISKQNVNQTFLKILGAILSAKLREKLKMLKQR	317
Qy	238	GPPPSFLELMKVIREEBEEAEAFENES	263

Blank

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 13:35:22 ; Search time 52.4349 Seconds
(without alignments)
1194.261 Million cell updates/sec

Title: US-10-037-860-7
Perfect score: 996
Sequence: 1 PLALDWCRTMSVDEQKSL.....EESFEVLEQATEIVKWP 195

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1342398 seqs, 321133274 residues

Total number of hits satisfying chosen parameters: 1342398

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	996	100.0	125	13	US-10-037-860-7
2	604	60.6	120	12	US-09-804-014A-42
3	593	59.5	283	13	US-10-037-860-11
4	466	46.8	353	9	US-09-965-529-7
5	466	46.8	353	10	US-09-969-680A-7
6	462.5	46.4	463	13	US-10-037-860-13
7	441	44.3	329	13	US-10-037-860-4
8	440	44.2	318	12	US-09-804-014A-40
9	436	43.8	312	12	US-09-804-014A-73
10	436	43.8	312	12	US-09-804-014A-74
11	436	43.8	321	12	US-09-804-014A-39
12	436	43.8	351	9	US-09-965-529-1
13	436	43.8	351	10	US-09-969-680A-1
14	436	43.8	351	12	US-09-804-014A-16
15	436	43.8	351	15	US-10-341-434-10

16	434	43.6	452	16	US-10-408-765A-2385
17	345	34.6	399	15	US-10-094-749-1978
18	327.5	32.9	204	14	US-10-029-386-33747
19	306	30.7	120	12	US-09-804-014A-41
20	270	27.1	116	9	US-09-864-761-34645
21	146	14.7	538	16	US-10-408-765A-2392
22	137.5	13.8	584	12	US-10-221-278-355
23	137.5	13.8	584	15	US-10-291-172-355
24	99.5	10.0	403	15	US-10-094-466-38
25	92.5	9.3	337	12	US-10-296-115-1208
26	89	8.9	342	13	US-10-001-857-201
27	87.5	8.8	255	12	US-10-087-192-213
28	87.5	8.8	311	12	US-09-727-100-1
29	87.5	8.8	1357	15	US-10-295-027-1199
30	87.5	8.8	3859	16	US-10-408-765A-354
31	86	8.6	1083	15	US-10-369-493-4443
32	86	8.6	1083	15	US-10-369-493-7202
33	86	8.6	1084	12	US-10-282-122A-49912
34	85.5	8.6	407	15	US-10-369-493-17903
35	84.5	8.5	520	9	US-09-213-678-2
36	84.5	8.5	520	14	US-10-032-585-7035
37	84	8.4	336	9	US-09-745-763-17
38	84	8.4	1638	13	US-10-090-458-2
39	84	8.4	1642	13	US-10-090-458-5
40	84	8.4	1642	14	US-10-005-338B-5
41	84	8.4	3613	14	US-10-156-761-10432
42	83.5	8.4	492	9	US-09-764-898-206
43	82	8.2	408	12	US-10-425-114-55513
44	82	8.2	547	16	US-10-437-963-199553
45	82	8.2	607	12	US-10-424-599-218036

ALIGNMENTS

RESULT 1

US-10-037-860-7
; Sequence 7, Application US/10037860
; Publication No. US20020123114A1
; GENERAL INFORMATION:
; APPLICANT: Jerome B. Posner
; APPLICANT: Joseph O. Dalmat
; APPLICANT: Myrina R. Rosenfeld
; TITLE OF INVENTION: Ma FAMILY POLYPEPTIDES AND ANTI-Ma
; TITLE OF INVENTION: ANTIBODIES
; FILE REFERENCE: 2581.1004-004
; CURRENT APPLICATION NUMBER: US/10/037,860
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 09/189,527
; PRIOR FILING DATE: 1998-11-10
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 195
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-037-860-7

Query Match 100.0%; Score 996; DB 13; Length 195;
Best Local Similarity 100.0%; Pred. No. 2.8e-95;
Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	PLALDWCRTMSVDEQKSLMTGIPADFEAEIQEVLQETLKSIGRYRLGKIFRQEN	60
Db	1	PLALDWCRTMSVDEQKSLMTGIPADFEAEIQEVLQETLKSIGRYRLGKIFRQEN	60
Qy	61	ANAVLLELLEDVSAIPSEVQGGKGVKVFKTPNQDTFELRNLNLFLEKGTQSGMF	120
Db	61	ANAVLLELLEDVSAIPSEVQGGKGVKVFKTPNQDTFELRNLNLFLEKGTQSGMF	120
Qy	121	RALQGEALSPATVPCISPELLAHLLGQMAHAPOLPLPMRYKLRVFGSVAPEEESF	180
Db	121	RALQGEALSPATVPCISPELLAHLLGQMAHAPOLPLPMRYKLRVFGSVAPEEESF	180

181 EVWLEQATEIVKEWP 195
|||||
181 EVWLEQATEIVKEWP 195

RESULT 2
US-09-804-014A-42
; Sequence 42, Application US/09804014A
; Publication No. US20030064489A1
; GENERAL INFORMATION:
; APPLICANT: Li, Li
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Vernet, Corine
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard
; APPLICANT: Spaderna, Steven
; APPLICANT: Majumder, Kumud
; TITLE OF INVENTION: Novel Polyvinylides and Nucleic Acids Encoding Same

```

; APPLICANT: Majumder, Kumud
; TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-721 US
; CURRENT APPLICATION NUMBER: US/09/804,014A
; CURRENT FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: 60/188,316
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 60/188,277
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 60/189,139
; PRIOR FILING DATE: 2000-03-14
; PRIOR APPLICATION NUMBER: 60/189,140
; PRIOR FILING DATE: 2000-03-14
; PRIOR APPLICATION NUMBER: 60/190,401
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/190,231
; PRIOR FILING DATE: 2000-03-17
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 42
; LENGTH: 120
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-804-014A-42

```

[illegible]

RESULT 5
US-09-969-680A-7
: Sequence 7, Application US/09969680A


```
; APPLICANT: Shimkets, Richard
; APPLICANT: Spaderna, Steven
; APPLICANT: Majumder, Kumud
; TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-721 US
; CURRENT APPLICATION NUMBER: US/09/804,014A
; CURRENT FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: 60/188,316
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 60/188,277
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 60/189,139
; PRIOR FILING DATE: 2000-03-14
; PRIOR APPLICATION NUMBER: 60/189,140
; PRIOR FILING DATE: 2000-03-14
; PRIOR APPLICATION NUMBER: 60/190,401
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/190,231
; PRIOR FILING DATE: 2000-03-17
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 40
; LENGTH: 318
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (20)
; OTHER INFORMATION: Wherein Xaa is any amino acid as defined in the
; OTHER INFORMATION: specification
US-09-804-014A-40

Query Match          44.2%; Score 440; DB 12; Length 318;
Best Local Similarity 46.4%; Pred. No. 4.7e-37;
Matches 90; Conservative 36; Mismatches 62; Indels 6; Gaps 4;

Qy      2 LALLDWCRI MSVDQSLSMTGTPADPFEAEIOEVLOETLKSIGRYELLGKIFPKQENA 61
Db      : |||||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Qy      3 MTLLEDWCRGM DVNSQRXLLVWGIPVNCDEAEIETLQAAMPQVS-YMILGRMPFRENA 61
Db      : |||||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Qy      62 NAVLELLEDTDVAISAEIVSGVGKGVKVIKPTNQDTFLERLNLFLEKEGQTVSGMFR 121
Db      : |||||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Qy      62 KAALLELTGAVDYAAIPREMPCKGGVKVLFPKPTSDAEFLERHLHFLAREGTVDVAR 121
Db      : |||||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Qy      122 ALGOEALSPATVPVCISPSELLAHLGQAHAAPQLPL-PMRVKKURVFGSVAVPPEESF 180
Db      : |||||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Qy      122 VLGFQ--NPTPT--GPMPAEMLYLDNVLIQPLVESIWYKRITLFSCKGHPRAWRGNF 177
Db      : |||||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Qy      181 EVWLEQATEIVKEW 194
Db      : |||||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Qy      178 DPWLEHTNEVLEE 191
Db      : |||||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

RESULT 9
US-09-804-014A-73
; Sequence 73, Application US/09804014A
; Publication No. US20030064489A1
; GENERAL INFORMATION:
; APPLICANT: Li, Li
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Vernet, Corine
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard
; APPLICANT: Spaderna, Steven
; APPLICANT: Majumder, Kumud
; TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-721 US
; CURRENT APPLICATION NUMBER: US/09/804,014A
; CURRENT FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: 60/188,316
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 60/188,277
; PRIOR FILING DATE: 2000-03-10
```



```

Query Match      43.8%; Score 436; DB 9; Length 351;
Best Local Similarity 46.9%; Pred. No. 1.4e-36;
Matches 91; Conservative 30; Mismatches 71; Indels 2; Gaps 2

QY      2 LALLEDCWRIMSVDEOKSLMVTGIPADFEAEAEIQEVLTQETIKSLGRYRLLLGKIFRKQENA 61
DB      3 LRLLEDCWRGMDWMPKRALHGIAGISQCSVAETEEALQAGLAPLGEYRLLLGRMFRFRDENR 62

QY      62 NAVILLELDTVSALPSEVQKGGVWKVFKTPNQDTEFLERLNLFLKEGQTVSGMFR 121
DB      63 KVALVGLTAETSHALVPKEIPGKGIWRVIFPKPPDPDNTFLSLNEFLAGEMTVUGLSR 122

QY      122 ALQGEALS PATVPCISPELLAHLHGQMAHAPQPLP-MRYRKLRVFGSGSAVPAPEBSF 180
DB      123 ALCHENGSLDPEQGMTPENWAPMLAQAL-EALQALQCLYKVKLRVFGSGSPPEGEEF 181

QY      181 EYWLEQATEIVKEW 194
DB      182 GRWMFHTTQIMKAW 195

RESULT 13
US-09-969-680A-1
; Sequence 1, Application US/09969680A
; Publication No. US20030124649A1
; GENERAL INFORMATION:
; APPLICANT: LAL, Preeti; YUB, Henry
; APPLICANT: TANG, Y. Tom; BANDMAN, Olga
; APPLICANT: BURFORD, Neil; AZIMZAI, Yalda
; APPLICANT: BAUGHN, Mariah R.; LU, Pyung Aina M.
; APPLICANT: PATTERSON, Chandra
; TITLE OF INVENTION: MEMBRANE ASSOCIATED PROTEINS
; FILE REFERENCE: PF-0731-1 USA
; CURRENT APPLICATION NUMBER: US/09/969,680A
; CURRENT FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US00/22315
; PRIOR FILING DATE: 2000-08-14

```

Prior Application Number: 60/149,641
Prior Filing Date: 1999-08-17
Prior Application Number: 60/164,203
Prior Filing Date: 1999-11-09
Number of Seq ID NOS: 74
Software: PERL Program
Seq ID No 1
Length: 351
Type: PRT
Organism: Homo sapiens
Feature: misc feature
Name/Key: Inocyte ID No. US20030124649A1 112301CD1
US-09-969-680A-1

Query Match 43.8%; Score 436; DB 10; Length 351;
Best Local Similarity 46.9%; Pred. No. 1.4e-36;
Matches 91; Conservative 30; Mismatches 71; Indels 2; Gaps 2;

QY 2 LALLEDCRIMSVDEQKSLMVTGIPADFEAEIOEVLQETLKSIGRYRLGKIPRKQENA 61
Db 3 LRLLDWCRCMDMNPRAKLLIAGISQCSVAIEEALQAGLAPLGEYRLGMRFRDENR 62

QY 62 NAVILLELDDTDSVAIPSEVQGGVWVKVIFKTPNQDTEFLERLNLFLKEGQTVSGMFR 121
Db 63 KVALVGLTAETSHALVPKEIPGKGIWRVIFKPPDPNTFLSRLNEFLAGEGTVGELSR 122

QY 122 ALGQEALSPATVPCISPELLAHLGQAMAHAPOLLP-MRYRKLRFVSGSAVPAPEESF 180
Db 123 ALGHENGLDPEQGMIPEMWAPMLAQAAL-EALQALQCLKYKKLRVFSGRESPEPEGE 181

QY 181 EVWLEQATEIVKEW 194
Db 182 GRWMFHTTQMIKAW 195

RESULT 15
US-10-341-434-10
Sequence 10, Application US/10341434
Publication No. US20030215835A1
GENERAL INFORMATION:
APPLICANT: Origene Technologies
TITLE OF INVENTION: Differentially Regulated Prostate Cancer Genes
FILE REFERENCE: 9U 204 205 R1
CURRENT APPLICATION NUMBER: US/10/341,434
CURRENT FILING DATE: 2003-07-18
PRIOR APPLICATION NUMBER: US 60/348,164
PRIOR FILING DATE: 2002-01-15
PRIOR APPLICATION NUMBER: US 60/348,119
PRIOR FILING DATE: 2002-01-15
NUMBER OF SEQ ID NOS: 238
SOFTWARE: PatentIn version 3.1
SEQ ID NO 10
LENGTH: 351
TYPE: PRT
ORGANISM: Homo sapiens
US-10-341-434-10

Query Match 43.8%; Score 436; DB 15; Length 351;
Best Local Similarity 46.9%; Pred. No. 1.4e-36;
Matches 91; Conservative 30; Mismatches 71; Indels 2; Gaps 2;

QY 2 LALLEDCRIMSVDEQKSLMVTGIPADFEAEIOEVLQETLKSIGRYRLGKIPRKQENA 61
Db 3 LRLLDWCRCMDMNPRAKLLIAGISQCSVAIEEALQAGLAPLGEYRLGMRFRDENR 62

QY 62 NAVILLELDDTDSVAIPSEVQGGVWVKVIFKTPNQDTEFLERLNLFLKEGQTVSGMFR 121
Db 63 KVALVGLTAETSHALVPKEIPGKGIWRVIFKPPDPNTFLSRLNEFLAGEGTVGELSR 122

QY 122 ALGQEALSPATVPCISPELLAHLGQAMAHAPOLLP-MRYRKLRFVSGSAVPAPEESF 180
Db 123 ALGHENGLDPEQGMIPEMWAPMLAQAAL-EALQALQCLKYKKLRVFSGRESPEPEGE 181

QY 181 EVWLEQATEIVKEW 194
Db 182 GRWMFHTTQMIKAW 195

Search completed: September 21, 2004, 14:04:07
Job time : 53.4349 secs

Prior Application Number: 60/149,641
Prior Filing Date: 1999-08-17
Prior Application Number: 60/164,203
Prior Filing Date: 1999-11-09
Number of Seq ID NOS: 74
Software: PERL Program
Seq ID No 1
Length: 351
Type: PRT
Organism: Homo sapiens
Feature: misc feature
Name/Key: Inocyte ID No. US20030124649A1 112301CD1
US-09-969-680A-1

Query Match 43.8%; Score 436; DB 10; Length 351;
Best Local Similarity 46.9%; Pred. No. 1.4e-36;
Matches 91; Conservative 30; Mismatches 71; Indels 2; Gaps 2;

QY 2 LALLEDCRIMSVDEQKSLMVTGIPADFEAEIOEVLQETLKSIGRYRLGKIPRKQENA 61
Db 3 LRLLDWCRCMDMNPRAKLLIAGISQCSVAIEEALQAGLAPLGEYRLGMRFRDENR 62

QY 62 NAVILLELDDTDSVAIPSEVQGGVWVKVIFKTPNQDTEFLERLNLFLKEGQTVSGMFR 121
Db 63 KVALVGLTAETSHALVPKEIPGKGIWRVIFKPPDPNTFLSRLNEFLAGEGTVGELSR 122

QY 122 ALGQEALSPATVPCISPELLAHLGQAMAHAPOLLP-MRYRKLRFVSGSAVPAPEESF 180
Db 123 ALGHENGLDPEQGMIPEMWAPMLAQAAL-EALQALQCLKYKKLRVFSGRESPEPEGE 181

QY 181 EVWLEQATEIVKEW 194
Db 182 GRWMFHTTQMIKAW 195

RESULT 14
US-09-804-014A-16
Sequence 16, Application US/09804014A
Publication No. US20030064489A1
GENERAL INFORMATION:
APPLICANT: Li, Li
APPLICANT: Padigaru, Muralidhara
APPLICANT: Vernet, Corine
APPLICANT: Fernandes, Elma
APPLICANT: Shinkets, Richard
APPLICANT: Spaderna, Steven
TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoding Same
FILE REFERENCE: 15966-721 US
CURRENT APPLICATION NUMBER: US/09/804,014A
CURRENT FILING DATE: 2002-04-24
PRIOR APPLICATION NUMBER: 60/188,316
PRIOR FILING DATE: 2000-03-10
PRIOR APPLICATION NUMBER: 60/188,277
PRIOR FILING DATE: 2000-03-10
PRIOR APPLICATION NUMBER: 60/189,139
PRIOR FILING DATE: 2000-03-14
PRIOR APPLICATION NUMBER: 60/189,140
PRIOR FILING DATE: 2000-03-14
PRIOR APPLICATION NUMBER: 60/190,401
PRIOR FILING DATE: 2000-03-17
PRIOR APPLICATION NUMBER: 60/190,231
PRIOR FILING DATE: 2000-03-17
NUMBER OF SEQ ID NOS: 75
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 16
LENGTH: 351
TYPE: PRT
ORGANISM: Homo sapiens
US-09-804-014A-16

Query Match 43.8%; Score 436; DB 12; Length 351;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 13:20:21 ; Search time 36.0037 Seconds
(without alignments)
1530.308 Million cell updates/sec

Title: US-10-037-860-7

Perfect score: 996

Sequence: 1 PLALLEDWCRIMSVDEQKSL.....EESEFVWLEQATEIVKEWP 195

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04.*

- 1: Geneseqp1980s.*
- 2: Geneseqp1990s.*
- 3: Geneseqp2000s.*
- 4: Geneseqp2001s.*
- 5: Geneseqp2002s.*
- 6: Geneseqp2003as.*
- 7: Geneseqp2003bs.*
- 8: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	996	100.0	195	3	AAB12526	Aab12526 Human Ma2
2	982	98.6	364	7	ADC08977	Adc08977 Onconeuro
3	590	59.2	283	3	AAB12528	Aab12528 Human Ma4
4	466	46.8	353	4	AAB74701	Aab74701 Human mem
5	462.5	46.4	455	5	ABB05727	Abb05727 Human s19
6	462.5	46.4	463	3	AAB12529	Aab12529 Human Ma5
7	462.5	46.4	463	3	AAB42315	Aab42315 Human ORF
8	436	43.8	312	3	AAB43023	Aab43023 Human ORF
9	436	43.8	351	4	AAE01340	Aae01340 Human gen
10	436	43.8	351	4	AAB74695	Aab74695 Human mem
11	436	43.8	351	4	AAU08664	Aau08664 Human NOV
12	435	43.7	329	3	AAB12525	Aab12525 Human Ma1
13	434	43.6	452	6	AAO16179	Aao16179 Human pro
14	345	34.6	399	6	ADA54410	Ada54410 Human pro
15	345	34.6	399	6	ABG99947	Abg99947 Human nov
16	342	34.3	439	6	ABP75736	Abp75736 Human sec
17	270	27.1	116	4	AAm14937	Aam14937 Peptide #
18	270	27.1	116	4	ABB33906	Abb33906 Peptide #
19	270	27.1	116	4	AAm27367	Aam27367 Peptide #
20	270	27.1	116	4	ABB28727	Abb28727 Peptide #
21	270	27.1	116	4	ABB19347	Abb19347 Protein #
22	270	27.1	116	4	AAm67075	Aam67075 Human bon
23	270	27.1	116	4	AAm54673	Aam54673 Human bra
24	270	27.1	116	4	ABg48741	Abg48741 Human liv
25	270	27.1	116	4	AAm02665	Aam02665 Peptide #

ALIGNMENTS

RESULT 1
AAB12526
ID AAB12526 standard; protein; 195 AA.
XX
AC AAB12526;
XX
DT 02-NOV-2000 (first entry)
XX
DE Human Ma2 protein SEQ ID NO:7.
XX
KW Ma1; Ma2; Ma3; Ma4; Ma5; paraneoplastic limbic encephalitis; diagnosis;
KW paraneoplastic syndrome; neoplasm; colon cancer; brainstem encephalitis;
KW breast cancer; parotid gland cancer; lung cancer; testicular cancer;
KW germ-cell tumour.
XX
OS Homo sapiens.
XX
PN JP2000146982-A.
XX
PD 26-MAY-2000.
XX
PF 10-NOV-1999; 99JP-00320171.
XX
PR 10-NOV-1998; 98US-00189527.
XX
PA (SLOK) SLOAN KETTERING INST CANCER RES.
XX
DR WPI; 2000-468119/41.
XX
N-PSDB; AAA60834.
XX
PT Diagnosing paraneoplastic syndrome e.g. paraneoplastic limbic
PT encephalitis or neoplasm e.g. colon cancer comprising assessing a test
PT sample for the presence or absence of antibodies to a Ma family
PT polypeptide.
XX
PS Claim 48; Fig 2; 27pp; Japanese.
XX
CC The present invention describes a method for diagnosing a paraneoplastic
CC syndrome or neoplasm. The method comprises assessing a test sample for
CC the presence or absence of antibodies to a Ma family polypeptide (I). The
CC method is used to diagnose a paraneoplastic syndrome especially
CC paraneoplastic limbic encephalitis and/or brainstem encephalitis or
CC neoplasm in an individual. The method diagnoses the neoplasm by assessing
CC antibodies to (i) preferably Ma1, which is indicative presence of breast
CC cancer, colon cancer, parotid gland cancer, lung cancer, testicular
CC cancer, germ-cell tumours or Ma2, which is indicative of testicular
CC cancer, germ-cell tumour, and lung cancer. The present sequence is the
CC Ma2 protein as given in the present invention

Abg36734 Human pep
Aae01336 Human gen
Abu52641 Human bra
Adc31124 Human nov
Abg19651 Novel hum
Abg14259 Novel hum
Aab94854 Human pro
Aau28186 Novel hum
Abg97495 Human NOV
Aam51624 KIAA0883 -
Abo14772 Novel hum
Abo14773 Novel hum
Aab60478 Human cel
Aam55693 Human pro
Abg13816 Novel hum
Ade08994 Novel pro
Abg12037 Novel hum
Abg05411 Novel hum
Abp52958 Human lun
Aar43684 Lipopolys

26 270 27.1 116 5 ABG36734
27 160 16.1 280 4 AAE01336
28 146 14.7 538 4 ABU52641
29 146 14.7 538 7 ADC31124
30 144 14.5 615 4 ABG19651
31 144 14.5 615 4 ABG14259
32 142 14.3 237 4 AAB94854
33 137.5 13.8 584 4 AAU28186
34 99.5 10.0 403 5 ABG97495
35 99.5 10.0 403 5 AAM51624
36 99.5 10.0 403 6 ABO14772
37 99.5 10.0 403 6 ABO14773
38 97.5 9.8 402 4 AAB60478
39 92.5 9.3 337 4 AAM55693
40 92 9.2 872 4 ABG13816
41 92 9.2 872 7 ADE08994
42 90 9.0 521 4 ABG12037
43 90 9.0 521 4 ABG05411
44 89 8.9 342 5 ABP52958
45 87.5 8.8 311 2 AAR43684

XX SQ Sequence 195 AA;
 Query Match 100.0%; Score 996; DB 3; Length 195;
 Best Local Similarity 100.0%; Pred. NO. 2.1e-96;
 Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLALLEDCRIMSVDEQKSLMTGTPADFEAEIQEVLQETLKSGLRYLLGKIFRKQEN 60
 DB 1 PLALLEDCRIMSVDEQKSLMTGTPADFEAEIQEVLQETLKSGLRYLLGKIFRKQEN 60

QY 61 ANAVLLELLEDDTDSVAIPSEVQKGGVWVKVIFKTPNQDTFELRLNLFLEKEGQTVSGMP 120
 DB 61 ANAVLLELLEDDTDSVAIPSEVQKGGVWVKVIFKTPNQDTFELRLNLFLEKEGQTVSGMF 120

QY 121 RALGQEALSPATVPCISPELLAHLGQAMAHAPQPLLPMRYKLRVFGSSAVPAPEESF 180
 DB 121 RALGQEALSPATVPCISPELLAHLGQAMAHAPQPLLPMRYKLRVFGSSAVPAPEESF 180

QY 181 EVWLEQATEIVKEMP 195
 DB 181 EVWLEQATEIVKEMP 195

RESULT 2
 ADC08977
 ID ADC08977 standard; protein; 364 AA.
 AC ADC08977;
 DT 18-DEC-2003 (first entry)
 DE Onconeural antigen Ma2 protein.
 KW Human; Ma2; onconeural; antigen; Alzheimer's disease;
 KW neurodegenerative disease; diagnosis; neuroprotective; gene therapy.
 OS Homo sapiens.
 XX W02003073104-A2.
 XX 04-SEP-2003.
 XX 26-FEB-2003; 2003WO-EP001946.
 XX 26-FEB-2002; 2002EP-00004177.
 XX 26-FEB-2002; 2002US-0359307P.
 XX (EVOT-) EVOTEC NEUROSCIENCES GMBH.
 XX Hipfel R, Von Der Kammer H, Pohlner J;
 XX WPI; 2003-721818/68.
 XX GENBANK; O94959, KIAA0883.
 XX Diagnosing or prognosticating, or determining increased risk of
 PT developing a neurodegenerative disease by determining level or activity
 PT of a transcription or translation product of a gene coding for Ma
 PT onconeural antigen.
 XX Disclosure; Fig 9; 51pp; English.
 XX The present sequence is the protein sequence of human onconeural
 CC antigen Ma2. The invention discloses the detection and differential
 CC expression and regulation of the Ma2 gene in specific brain regions of AD
 CC patients. The Ma2 gene and its transcription and/or translation products
 CC may have a causative role in the regional selective neuronal degeneration
 CC typically observed in AD, or may confer a neuroprotective function to the
 CC remaining nerve cells. Methods are claimed for diagnosing or
 CC prognosticating a neurodegenerative disease, for monitoring the
 CC progression of a neurodegenerative disease, and for evaluating treatment
 CC of a neurodegenerative disease, especially AD, in a subject by
 CC determining the level and/or activity of a transcription or translation

CC product of an Ma onconeural antigen gene, especially Ma2. Also claimed
 CC are: a method for treating or preventing AD and related neurodegenerative
 CC disorders using the Ma2 gene or its transcription or translation product;
 CC a method of screening for modulating agents of neurodegenerative diseases
 CC ; and a recombinant non-human animal comprising an Ma2 gene sequence,
 CC which is useful for screening, testing and validating candidate
 CC diagnostic and therapeutic agents.
 XX Sequence 364 AA;
 SQ Query Match 98.6%; Score 982; DB 7; Length 364;
 Best Local Similarity 99.0%; Pred. NO. 1.1e-94;
 Matches 192; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LALLEDCRIMSVDEQKSLMTGTPADFEAEIQEVLQETLKSGLRYLLGKIFRKQENA 61
 DB 3 LALLEDCRIMSVDEQKSLMTGTPADFEAEIQEVLQETLKSGLRYLLGKIFRKQENA 62

QY 62 NAVLLELLEDDTDSVAIPSEVQKGGVWVKVIFKTPNQDTFELRLNLFLEKEGQTVSGMR 121
 DB 63 NAVLLELLEDDTDSVAIPSEVQKGGVWVKVIFKTPNQDTFELRLNLFLEKEGQTVSGMR 122

QY 122 ALGOEALSPATVPCISPELLAHLGQAMAHAPQPLLPMRYKLRVFGSSAVPAPEESFE 181
 DB 123 ALGOEGVSPATVPCISPELLAHLGQAMAHAPQPLLPMRYKLRVFGSSAVPAPEESFE 182

QY 182 VWLEQATEIVKEMP 195
 DB 183 VWLEQATEIVKEMP 196

RESULT 3
 AAB12528
 ID AAB12528 standard; protein; 283 AA.
 XX AAB12528;
 XX 02-NOV-2000 (first entry)
 DT Human Ma4 protein SEQ ID NO:11.
 DE Ma1; Ma2; Ma3; Ma4; Ma5; paraneoplastic limbic encephalitis; diagnosis;
 KW paraneoplastic syndrome; neoplasm; colon cancer; brainstem encephalitis;
 KW breast cancer; parotid gland cancer; lung cancer; testicular cancer;
 KW germ-cell tumour.
 XX Homo sapiens.
 XX JF2000146982-A.
 XX 26-MAY-2000.
 XX 10-NOV-1999; 99JP-00320171.
 XX 10-NOV-1998; 98US-00189527.
 XX (SLOK) SLOAN KETTERING INST CANCER RES.
 XX WPI; 2000-468119/41.
 XX N-PSDB; AAA60836.
 XX Diagnosing paraneoplastic syndrome e.g. paraneoplastic limbic
 PT encephalitis or neoplasm e.g. colon cancer comprising assessing a test
 PT sample for the presence or absence of antibodies to a Ma family
 PT polypeptide.
 XX Claim 48; Fig 7-8; 27pp; Japanese.
 XX The present invention describes a method for diagnosing a paraneoplastic
 CC syndrome or neoplasm. The method comprises assessing a test sample for
 CC the presence or absence of antibodies to a Ma family polypeptide (1). The
 CC method is used to diagnose a paraneoplastic syndrome especially
 CC paraneoplastic limbic encephalitis and/or brainstem encephalitis or

CC neoplasm in an individual. The method diagnoses the neoplasm by assessing
 CC antibodies to (I) preferably Ma1, which is indicative presence of breast
 CC cancer, colon cancer, parotid gland cancer, lung cancer, testicular
 CC cancer and germ-cell tumours or Ma2, which is indicative of testicular
 CC cancer, germ-cell tumour, and lung cancer. The present sequence is the
 CC Ma4 protein as given in the present invention
 XX
 SQ

Sequence 283 AA;

Query Match 59.2%; Score 590; DB 3; Length 283;
 Best Local Similarity 97.4%; Pred. No. 2.3e-53;
 Matches 112; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 81 VGGKGVKVIPTNQDTFELRLNLFLEKQGVSGMFRALGOEALSPATVPCISPEL 140

Db 1 VGGKGVKVIPTNQDTFELRLNLFLEKQGVSGMFRALGOEALSPATVPCISPEL 60

QY 141 LAHLGQAMAHAPQLLPMRYKRLRVFSGSAPVAPPEESFEVWLEQATEIVKEMP 195

Db 61 LAHLGQAMAHAPQLLPMRYKRLRVFSGSAPVAPPEEDSFEVWLEQATEIVKEMP 115

RESULT 4

AAB74701

ID AAB74701 standard; protein; 353 AA.

AC AAB74701;

XX

DT 12-JUN-2001 (first entry)

XX

DE Human membrane associated protein MEMAP-7.

XX

KW Human; membrane associated protein; MEMAP; diagnosis; cytostatic;
 KW antiinflammatory; anticonvulsant; immunosuppressive; antidiarrheic;
 KW antiarteriosclerotic; gene therapy; cell proliferative disorder;
 KW autoimmune disorder; inflammatory disorder; neurological disorder;
 KW gastrointestinal disorder; cancer; inflammation; atherosclerosis;
 KW epilepsy; diarrhoea.

XX

OS Homo sapiens.

XX

FN WO200112662-A2.

XX

PD 22-FEB-2001.

XX

PF 14-AUG-2000; 2000WO-US022315.

XX

PR 17-AUG-1999; 99US-0149641P.

XX

PR 09-NOV-1999; 99US-0164203P.

XX

PA (INCY-) INCYTE GENOMICS INC.

XX

PI Lal P, Yue H, Tang YT, Bandman O, Burford N, Azimzai Y;

XX

PI Baughn MR, Lu DAM, Patterson C;

XX

DR WPI; 2001-168860/17.

XX

DR N-PSDB; AAF81747.

XX

Isolated polypeptide with a human membrane associated protein sequence is

PT useful for the diagnosis, prevention and treatment of cell proliferative,

PT autoimmune/inflammatory, neurological and gastrointestinal disorders.

XX

PS Claim 1; Page 119-120; 173pp; English.

XX

AAF81741 to AAF81777 encode the human membrane associated proteins

CC (MEMAP) given in AAB74695 to AAB74731. MEMAPs have cytostatic,

CC antiinflammatory, anticonvulsant, immunosuppressive, antidiarrheic and

CC antiarteriosclerotic activities, which can be used in gene therapy.

CC MEMAPs and agonist of MEMAPs can be used to treat a disease or condition

CC associated with decreased expression of functional MEMAP and antagonists

CC of MEMAP are used to treat a disease or condition associated with

CC overexpression of functional MEMAP. These disorders include cell

CC proliferative, autoimmune/inflammatory, neurological and gastrointestinal

CC disorders. The MEMAP polynucleotides and proteins are also used for the
 CC diagnosis of these disorders. Specific examples of these disorders
 CC include cancer, inflammation, atherosclerosis, epilepsy and diarrhoea.
 CC MEMAP proteins can be used to screen for compounds which specifically
 CC bind MEMAP including antibodies, oligonucleotides, proteins and small
 CC molecules. MEMAP polynucleotides can be used to prepare transgenic
 CC animals which can be studied to provide information concerning human
 CC disease. Anti-MEMAP antibodies are useful in immunoassays for the
 CC detection of MEMAP protein and can be used as antagonists to treat or
 CC prevent a disorder associated with MEMAP. Polynucleotides encoding MEMAP
 CC can be delivered to target cells with genetic abnormalities with respect
 CC to the expression of MEMAP to treat or prevent a disorder associated with
 CC MEMAP

XX Sequence 353 AA;

Query Match 46.8%; Score 466; DB 4; Length 353;

Best Local Similarity 47.9%; Pred. No. 3.7e-40;

Matches 93; Conservative 38; Mismatches 57; Indels 6; Gaps 4;

QY 2 LALLEDCWCRIMSDVDEQKSLMTGTIPADFEAEIQEVLTLSLGRYLLGKIFRQENA 61

Db 3 MTLLEDWCGRGMDVNSQRALLVWGIPVNCDEABIEETLQAMPQVS-YRMLGRMFREENA 61

QY 62 NAVLLELLEDDTIVSAIPSEVOGKGVKVIPTNQDTFELRLNLFLEKQGVSGMFR 121

Db 62 KAALLELTCAVDYAAIPREMPGKGVKVIPTNQDTFELRLNLFLEKQGVSGMFR 121

QY 122 ALGOEALSPATVPCISPELLLAHLGQAMAHAPQLL-PMRYKRLRVFSGSAPVAPPEESF 180

Db 122 VLGFQ--NPTPT--GPMPAEMNLVILDNVIQPLVESIWYKRLTLFSGRDPGPEETF 177

QY 181 EVWLEQATEIVKEW 194

Db 178 DPWLEHTNEVLEEW 191

RESULT 5

ABB05727

ID ABB05727 standard; protein; 455 AA.

AC ABB05727;

XX

DT 30-APR-2002 (first entry)

XX

DE Human signal transduction protein clone tes3_5k22.

XX

KW Human; foetal brain; foetal kidney; melanoma; testis; amygdala;

XX gene therapy.

OS Homo sapiens.

XX

PN WO200198454-A2.

XX

PD 27-DEC-2001.

XX

PF 25-APR-2001; 2001WO-IB002050.

XX

PR 25-APR-2000; 2000US-0199380P.

XX

PA (GEHU-) GERMAN HUMAN GENOME PROJECT.

XX

PI Wiemann S;

XX

DR WPI; 2002-055860/07.

XX

DR N-PSDB; ABA93764.

XX

PT Human cDNA sequences and clones derived from human fetal brain, fetal

PT kidney, melanoma, testis and amygdala cDNA libraries, useful in genetic

PT screening and therapy.

XX

PS Claim 1; Page 382; 611pp; English.

XX

CC The present invention describes assemblages and computer readable media
 CC comprising novel human cDNA sequences and clones derived from human
 CC foetal brain, foetal kidney, melanoma, testis and anygdala cDNA
 CC libraries. ABA93702 to ABA93766 represent human cDNA sequences from the
 CC present invention which encode the proteins given in ABB05662 to
 CC ABB05729. The human cDNA sequences and clones can be used in gene
 CC therapy. The clones may be used in a variety of applications, for example
 CC they may be used in profiling assays, for providing large arrays of human
 CC genetic material for implementing large-scale screening strategies and
 CC for treating diseases via gene therapy procedures

XX
 SQ Sequence 455 AA;

Query Match 46.4%; Score 462.5; DB 5; Length 455;
 Best Local Similarity 47.7%; Pred. No. 1.3e-39;
 Matches 93; Conservative 34; Mismatches 65; Indels 3; Gaps 2;
 QY 1 PLALLEDWCRIMSVDFQKSLMTVGIPADFEABIQEVLQETLKSGLRYRLGKIFRKOEN 60
 Db 2 PLTLQDWCRCGEHLNTRCMLILGIPEDCGEDEFEEFTLQACRHLGRYRVIGRMFRREEN 61
 QY 61 ANAVILLELLEDDTVSAIPSEVQKGGVWVKVIFKTPNODTEFLERLNLFLKEGCTVSGMF 120
 Db 62 AQAILLELAQDIDYALLPREIPGKGPWEVIVKPRNSDGEFLNRLNRFLEERTVSDMN 121
 QY 121 RALGQEAALSPATVPCISPELLAHLGQAMAHAPQPLL-PMRYKRLRVFSGSAVPAPPEES 179
 Db 122 RVLGSDTNCAPRTVTSPEFWT--WAQTGLAAVQPLLEQMLYRELRFVSGNTISIPGALA 179
 QY 180 FEVWLEQATEIVKWEK 194
 Db 180 FDAWLEHTTEMLQMW 194

RESULT 6

AAB12529
 ID AAB12529 standard; protein; 463 AA.

XX
 AC AAB12529;

XX
 DT 02-NOV-2000 (first entry)

XX
 DE Human Ma5 protein SEQ ID NO:13.

XX Ma1; Ma2; Ma3; Ma4; Ma5; paraneoplastic limbic encephalitis; diagnosis;
 KW paraneoplastic syndrome; neoplasm; colon cancer; brainstem encephalitis;
 KW breast cancer; parotid gland cancer; lung cancer; testicular cancer;
 KW germ-cell tumour.

XX Homo sapiens.

XX JP2000146982-A.

XX
 PD 26-MAY-2000.

XX
 PF 10-NOV-1999; 99JP-00320171.

XX
 PR 10-NOV-1998; 98US-00189527.

XX (SLOK) SLOAN KETTERING INST CANCER RES.

XX
 DR WPI; 2000-468119/41.

XX
 DR N-PSDB; AAA60837.

XX Diagnosing paraneoplastic syndrome e.g. paraneoplastic limbic
 PT encephalitis or neoplasm e.g. colon cancer comprising assessing a test
 PT sample for the presence or absence of antibodies to a Ma family
 PT polypeptide.

XX Claim 48; Fig 9-10; 27pp; Japanese.

XX The present invention describes a method for diagnosing a paraneoplastic
 CC syndrome or neoplasm. The method comprises assessing a test sample for

CC the presence or absence of antibodies to a Ma family polypeptide (I). The
 CC method is used to diagnose a paraneoplastic syndrome especially
 CC paraneoplastic limbic encephalitis and/or brainstem encephalitis or
 CC neoplasm in an individual. The method diagnoses the neoplasm by assessing
 CC antibodies to (I) preferably Ma1, which is indicative presence of breast
 CC cancer, colon cancer, parotid gland cancer, lung cancer, testicular
 CC cancer, germ-cell tumours or Ma2, which is indicative of testicular
 CC cancer, germ-cell tumour, and lung cancer. The present sequence is the
 CC Ma5 protein as given in the present invention

XX
 SQ Sequence 463 AA;

Query Match 46.4%; Score 462.5; DB 3; Length 463;
 Best Local Similarity 47.7%; Pred. No. 1.3e-39;
 Matches 93; Conservative 34; Mismatches 65; Indels 3; Gaps 2;
 QY 1 PLALLEDWCRIMSVDFQKSLMTVGIPADFEABIQEVLQETLKSGLRYRLGKIFRKOEN 60
 Db 2 PLTLQDWCRCGEHLNTRCMLILGIPEDCGEDEFEEFTLQACRHLGRYRVIGRMFRREEN 61
 QY 61 ANAVILLELLEDDTVSAIPSEVQKGGVWVKVIFKTPNODTEFLERLNLFLKEGCTVSGMF 120
 Db 62 AQAILLELAQDIDYALLPREIPGKGPWEVIVKPRNSDGEFLNRLNRFLEERTVSDMN 121
 QY 121 RALGQEAALSPATVPCISPELLAHLGQAMAHAPQPLL-PMRYKRLRVFSGSAVPAPPEES 179
 Db 122 RVLGSDTNCAPRTVTSPEFWT--WAQTGLAAVQPLLEQMLYRELRFVSGNTISIPGALA 179
 QY 180 FEVWLEQATEIVKWEK 194
 Db 180 FDAWLEHTTEMLQMW 194

RESULT 7

AAB42315
 ID AAB42315 standard; protein; 463 AA.

XX
 AC AAB42315;

XX
 DT 08-FEB-2001 (first entry)

XX
 DE Human ORFX ORF2079 polypeptide sequence SEQ ID NO:4158.

XX Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
 KW vulnery; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
 KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
 KW hypotensive; dermatological; immunosuppressive; antineoplastic;
 KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
 KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 KW cholesterol ester storage; systemic lupus erythematosus; infection;
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
 KW thrombosis; contraceptive.

XX Homo sapiens.

XX
 PN WO200058473-A2.

XX
 PD 05-OCT-2000.

XX
 XX 31-MAR-2000; 2000WO-US008621.

XX
 XX 31-MAR-1999; 99US-0127607P.

XX
 PR 02-APR-1999; 99US-0127636P.

XX
 PR 05-APR-1999; 99US-0127728P.

XX
 PR 30-MAR-2000; 2000US-00540763.

XX (CURA-) CURAGEN CORP.

XX

PI Shimkets RA, Leach M;
 XX WPI; 2000-602362/57.
 DR N-PSDB; AAC76524.
 XX
 PT Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease.
 XX
 PS Claim 11; Page 3345-3347; 5507pp; English.
 XX
 CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytostatic; hepatotropic; vulnery;
 CC antiparasitic; antiparkinsonian; nontropic; neuroprotective; osteopathic;
 CC anticonvulsant; antiarthritic; immunosuppressant; immunostimulant;
 CC cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;
 CC dermatological; immunosuppressive; antiinflammatory; antibacterial;
 CC antiviral; antifungal; antirheumatic; antithyroid; and antianaemic. The
 CC sequences can be used for determining the presence of or predisposition
 CC to, or preventing or treating pathological conditions associated with an
 CC ORFX-associated disorder. The nucleic acids can be used to express ORFX
 CC proteins in gene therapy vectors. The proteins and nucleic acids may be
 CC used to treat cancers, proliferative disorders, neurodegenerative
 CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease,
 CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester
 CC storage, systemic lupus erythematosus, severe combined immunodeficiency
 CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune
 CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and
 CC cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to
 CC enhance coagulation; to inhibit thrombosis; and as a contraceptive
 XX
 SQ Sequence 463 AA;
 Query Match 46.4%; Score 462.5; DB 3; Length 463;
 Best Local Similarity 47.7%; Pred. No. 1.3e-39;
 Matches 93; Conservative 34; Mismatches 65; Indels 3; Gaps 2;
 QY 1 PLALLEDCRIMSDVDEQKSLMVTGTPADFEAEIOEVLTSLKSLGRVLLGKIFRQEN 60
 Db 2 PLTLQDWCRCHEHLNTRCMLILGIPEDCGDEFEETLQACRHLYRVRVTRGMRFRREN 61
 QY 61 ANAVLLELLEDDTDSAIPSEVQGGKGVKVPKTPNQDTEFLERLNLFLEKEGQTVSGMF 120
 Db 62 AQAILLELAQDIDVALLPREIPGKGPWEVIVKPNSDGDFNLNRLFRERRTVSDMN 121
 QY 121 RALQGEALSPATVPCISPELLAHLGQMAHAPQLLP-PMRYKRLRVFSGSAVPAPEES 179
 Db 122 RVLGSDTNCSPRVVTISPEFWT--WAQTLGAAVQPLLEQMLYRELVRVSGNTISIPGALA 179
 QY 180 FEVWLEQATEIVKEW 194
 Db 180 FDWLEHTTEMLQW 194

KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 KW cholesterol ester storage; systemic lupus erythematosus; infection;
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
 KW thrombosis; contraceptive.
 OS Homo sapiens.
 XX
 PN WO200058473-A2.
 XX
 PD 05-OCT-2000.
 XX
 PF 31-MAR-2000; 2000WO-US008621.
 XX
 PR 31-MAR-1999; 99US-0127607P.
 PR 02-APR-1999; 99US-0127636P.
 PR 05-APR-1999; 99US-0127728P.
 PR 30-MAR-2000; 2000US-00540763.
 XX
 PA (CUPA-) CURAGEN CORP.
 XX
 PI Shimkets RA, Leach M;
 XX
 DR WPI; 2000-602362/57.
 DR N-PSDB; AAC77232.
 XX
 PT Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease.
 XX
 PS Claim 11; Page 4759; 5507pp; English.
 XX
 CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytostatic; hepatotropic; vulnery;
 CC antiparasitic; antiparkinsonian; nontropic; neuroprotective; osteopathic;
 CC anticonvulsant; antiarthritic; immunosuppressant; immunostimulant;
 CC cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;
 CC dermatological; immunosuppressive; antiinflammatory; antibacterial;
 CC antiviral; antifungal; antirheumatic; antithyroid; and antianaemic. The
 CC sequences can be used for determining the presence of or predisposition
 CC to, or preventing or treating pathological conditions associated with an
 CC ORFX-associated disorder. The nucleic acids can be used to express ORFX
 CC proteins in gene therapy vectors. The proteins and nucleic acids may be
 CC used to treat cancers, proliferative disorders, neurodegenerative
 CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease,
 CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester
 CC storage, systemic lupus erythematosus, severe combined immunodeficiency
 CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune
 CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and
 CC cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to
 CC enhance coagulation; to inhibit thrombosis; and as a contraceptive
 XX
 SQ Sequence 312 AA;
 Query Match 43.8%; Score 436; DB 3; Length 312;
 Best Local Similarity 46.9%; Pred. No. 4.5e-37;
 Matches 91; Conservative 30; Mismatches 71; Indels 2; Gaps 2;
 QY 2 LALLEDCRIMSDVDEQKSLMVTGTPADFEAEIOEVLTSLKSLGRVLLGKIFRQENA 61
 Db 3 LRLLEDWCRCMDMNPRAKLLTAGISQSCSAEIEEALQAGLAPLGEVRLLRGMRFRREN 62
 QY 62 NAVLLELLEDDTDSAIPSEVQGGKGVKVPKTPNQDTEFLERLNLFLEKEGQTVSGMF 121
 Db 63 KVALVGLTAETSHALVPKEIPGKGIWRVIFKPPDPNTFLSLRNLNEFLAGEMTVGLSR 122
 QY 122 ALQGEALSPATVPCISPELLAHLGQMAHAPQLLP-MRYKRLRVFSGSAVPAPEES 180
 Db 123 ALGHENGLDPEQGMIPMWAPMLAQL-BALQALQCLKYKLRVFSGRESPEEGEEF 181
 QY 181 EVWLEQATEIVKEW 194

CC MEMAPs and agonist of MEMAPs can be used to treat a disease or condition
 CC associated with decreased expression of functional MEMAP and antagonists
 CC of MEMAP are used to treat a disease or condition associated with
 CC overexpression of functional MEMAP. These disorders include cell
 CC proliferative, autoimmune/inflammatory, neurological and gastrointestinal
 CC diseases. The MEMAP polynucleotides and proteins are also used for the
 CC diagnosis of these disorders. Specific examples of these disorders
 CC include cancer, inflammation, atherosclerosis, epilepsy and diarrhoea.
 CC MEMAP proteins can be used to screen for compounds which specifically
 CC bind MEMAP including antibodies, oligonucleotides, proteins and small
 CC molecules. MEMAP polynucleotides can be used to prepare transgenic
 CC animals which can be studied to provide information concerning human
 CC disease. Anti-MEMAP antibodies are useful in immunoassays for the
 CC detection of MEMAP protein and can be used as antagonists to treat or
 CC prevent a disorder associated with MEMAP. Polynucleotides encoding MEMAP
 CC can be delivered to target cells with genetic abnormalities with respect
 CC to the expression of MEMAP to treat or prevent a disorder associated with
 CC MEMAP
 XX
 XX Sequence 351 AA;

Query Match 43.8%; Score 436; DB 4; Length 351;
 Best Local Similarity 46.9%; Pred. No. 5.4e-37;
 Matches 91; Conservative 30; Mismatches 71; Indels 2; Gaps 2;
 QY 2 LALLEDWCRIMSVDEQKSLMTGIPADFEAEIQVLOETLKSGLRYLLGKIFRQENA 61
 DB ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
 3 LLELEDWCRGMDMNPFRKALLIAGISQSCVAIEEALQGLAPLGEYRLGMRFDENR 62
 QY 62 NAVLELLEDDTVSAIPSEVQGGVWVIFKTPNQDTFLERLNLFLEKQGVSGMFR 121
 DB || : : : : : || : : : : : || : : : : : || : : : : : || : : : : :
 63 KVALVGLTAETSHALVPKEIPCKGGIWRVIFKPPDPDNTFLSRLNEFLAGEMTVGELSR 122
 QY 122 ALGQFALSPATVPCISPELLAHLGQAMAHAPQLLP-MRYRKLRFVSGSAVPAPPEESF 180
 DB ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
 123 ALGHENGSLDPEQGMIPMWAPMLAQAL-EALQPALQCLKYKLRVFSGRSPPEGEF 181
 QY 181 EWLQATEIVKEM 194
 DB || : : : : : || : : : : : || : : : : : || : : : : : || : : : : :
 182 GRWMEHTTQMIKAW 195

RESULT 11
 AAU08664
 ID AAU08664 standard; protein; 351 AA.
 XX
 AC AAU08664;
 DT 18-DEC-2001 (first entry)
 XX
 DE Human NOV8 protein.
 XX
 KW Human; NOV8; cytostatic; neurotropic; neuroprotective; vulnary;
 KW cerebroprotective; antiparkinsonian; hypotensive; antiasthmatic;
 KW antidabetic; antiparotatic; antiinflammatory; immunosuppressive;
 KW antiatherosclerotic; dermatological; cancer; neurological disorder;
 KW parkinson's disease; diabetes mellitus; asthma; enamel defect;
 KW immune disorder; autoimmune disease; respiratory disorder; bone disorder;
 KW musculoskeletal disorders; leukaemia; lymphoma;
 KW cell growth regulation disorder; lesional psoriatic skin;
 KW atherosclerosis; abdominal aortic aneurysm.
 XX
 OS Homo sapiens.
 XX
 FN WO200168851-A2.
 XX
 PD 20-SEP-2001.
 XX
 PF 12-MAR-2001; 2001WO-US007735.
 XX
 PR 10-MAR-2000; 2000US-0189277P.
 PR 10-MAR-2000; 2000US-0189316P.
 PR 14-MAR-2000; 2000US-0189139P.

PR 14-MAR-2000; 2000US-0189140P.
 PR 17-MAR-2000; 2000US-0190231P.
 PR 17-MAR-2000; 2000US-0190401P.
 XX
 XX (CURA-) CURAGEN CORP.
 PA Padigaru M, Vernet CAM, Fernandes E, Shinkets RA, Spaderna SK;
 PI Majumder K, Li L;
 XX WPI; 2001-570869/64.
 DR N-PSDB; AAS13342.
 XX
 PT Novel polypeptides and nucleic acids homologous to members of collagen,
 PT potassium channel, tuftelin family of proteins for diagnosing, treating
 PT cancer, atherosclerosis, neurological, skin and enamel defect disorders.
 XX
 PS Claim 1; Page 29; 128pp; English.
 XX

CC The invention relates to isolated NOVX (NOVX1-11) polypeptides and the
 CC polynucleotides that encode them. NOVX polypeptides, polynucleotides and
 CC anti-NOVX antibodies are useful for treating or preventing a pathology
 CC associated with NOVX polypeptide in humans and for treating a syndrome
 CC associated with human disease e.g. disorders characterised by altered
 CC cell motility, proliferation and migration e.g. cancer, angiogenesis and
 CC wound healing (NOV1-3), neurological disorders, e.g. episodic ataxia,
 CC autosomal dominant myokymia, stroke, Parkinson's disease, Alzheimer's
 CC disease, non-insulin dependent diabetes mellitus, asthma, hypertension
 CC and seizure (NOVA), enamel defects, such as amelogenesis imperfecta and
 CC disorders involving enamel defects, including hypoplasia and
 CC hypomineralisation (NOV5-7), paraneoplastic neurological disorders, e.g.
 CC paraneoplastic limbic of brain-stem encephalitis occurring during
 CC testicular cancer, diabetes, reproductive health, metabolic and autoimmune
 CC disorders, gastrointestinal disorders, bone disorders, musculoskeletal
 CC disorders, respiratory/lymphoma and tissue/cell growth regulation disorders
 CC (NOV8), lesional psoriatic skin (NOV9-10) and atherosclerosis, abdominal
 CC aortic aneurysm and neurological disorders (NOV11). NOVX polypeptide is
 CC also useful for identifying an agent that binds to it and a cell
 CC expressing NOVX polypeptide is useful for identifying a therapeutic agent
 CC for use in treatment of a NOVX related pathology. The antibodies and a
 CC polypeptide having 95% sequence identity to NOVX polypeptide are useful
 CC for treating a pathological state in a mammal. The present sequence
 CC represents NOV8, a possible neuronal antigen-like protein
 XX
 SQ Sequence 351 AA;

Query Match 43.8%; Score 436; DB 4; Length 351;
 Best Local Similarity 46.9%; Pred. No. 5.4e-37;
 Matches 91; Conservative 30; Mismatches 71; Indels 2; Gaps 2;
 QY 2 LALLEDWCRIMSVDEQKSLMTGIPADFEAEIQVLOETLKSGLRYLLGKIFRQENA 61
 DB ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
 3 LLELEDWCRGMDMNPFRKALLIAGISQSCVAIEEALQGLAPLGEYRLGMRFDENR 62
 QY 62 NAVLELLEDDTVSAIPSEVQGGVWVIFKTPNQDTFLERLNLFLEKQGVSGMFR 121
 DB || : : : : : || : : : : : || : : : : : || : : : : : || : : : : :
 63 KVALVGLTAETSHALVPKEIPCKGGIWRVIFKPPDPDNTFLSRLNEFLAGEMTVGELSR 122
 QY 122 ALGQFALSPATVPCISPELLAHLGQAMAHAPQLLP-MRYRKLRFVSGSAVPAPPEESF 180
 DB ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
 123 ALGHENGSLDPEQGMIPMWAPMLAQAL-EALQPALQCLKYKLRVFSGRSPPEGEF 181
 QY 181 EWLQATEIVKEM 194
 DB || : : : : : || : : : : : || : : : : : || : : : : : || : : : : :
 182 GRWMEHTTQMIKAW 195

RESULT 12
 AAB12525
 ID AAB12525 standard; protein; 329 AA.
 XX
 AC AAB12525;
 XX

```
DT 02-NOV-2000 (first entry)
XX Human Mal protein SEQ ID NO:4.
DE
XX
XX
XX Ma1; Ma2; Ma3; Ma4; Ma5; paraneoplastic limbic encephalitis; diagnosis;
KW paraneoplastic syndrome; neoplasm; colon cancer; brainstem encephalitis;
KW breast cancer; parotid gland cancer; lung cancer; testicular cancer;
KW germ-cell tumour.
XX
XX Homo sapiens.
OS
XX
XX JP2000146982-A.
PN
XX
XX 26-MAY-2000.
PD
XX
XX 10-NOV-1999; 99JP-00320171.
PF
XX
XX 10-NOV-1998; 98US-00189527.
PR
XX
XX (SLOK ) SLOAN KETTERING INST CANCER RES.
PA
XX
XX WPI; 2000-468119/41.
DR N-PSDB; AAA60833.
XX
XX Diagnosing paraneoplastic syndrome e.g. paraneoplastic limbic
PT encephalitis or neoplasm e.g. colon cancer comprising assessing a test
PT sample for the presence or absence of antibodies to a Ma family
PT polypeptide.
XX
XX Claim 48; Fig 1; 27pp; Japanese.
PS
XX
XX The present invention describes a method for diagnosing a paraneoplastic
CC syndrome or neoplasm. The method comprises assessing a test sample for
CC the presence or absence of antibodies to a Ma family polypeptide (I). The
CC method is used to diagnose a paraneoplastic syndrome especially
CC paraneoplastic limbic encephalitis and/or brainstem encephalitis or
CC neoplasm in an individual. The method diagnoses the neoplasm by assessing
CC antibodies to (I) preferably Ma1, which is indicative presence of breast
CC cancer, colon cancer, parotid gland cancer, lung cancer, testicular
CC cancer and germ-cell tumours or Ma2, which is indicative of testicular
CC cancer, germ-cell tumour, and lung cancer. The present sequence is the
CC Ma1 protein as given in the present invention
XX
XX Sequence 329 AA;
SQ
Query Match 43.7%; Score 435; DB 3; Length 329;
Best Local Similarity 45.9%; Pred. No. 6.2e-37; Indels 6; Gaps 4;
Matches 89; Conservative 37; Mismatches 62;
QY 2 LALLEDCWCRIMSDVEQKSLMVTGIPADFEAEIOEVLQETLKSIGRYRLGKIFRKQENA 61
Db 3 MTLLEDWCRCMDVNSQRTLLVWGIPVNCDEAEIETLQAAMPQVS-YRMLGRMFWRENA 61
QY 62 NAVILLELLETDVSAIPSEVQGGKGVWVKVIFKTPNQDTEFLERINLFLEKEGQTVSGMFR 121
Db 62 KAALLETGTGAVDVAATPREMPGKGVWVKVLFKPTSDAEFLERLHLHLLAREGWTVDVAR 121
QY 122 ALGQEALSPATVPCI-SPELLAHLIGQAMAHAPQELL-PMRYKLRVFGSGAVPAPERESF 180
Db 122 VLGFQ--NPTPT--GPEMPAEMLNLYLDNVIQPLVESIWKYKTLFLSGKHPRWNGNF 177
QY 181 EVMLEQATEIVKKEW 194
Db 178 DPWLEHTNEVLEEW 191
RESULT 13
AAO16179
ID AAO16179 standard; protein; 452 AA.
XX
AC AAO16179;
XX
XX 28-MAR-2003 (first entry)
DT
```

```
XX Human protein #5.
DE
XX
XX Human; vaccine; adult whole brain; foetal whole brain; tonsil;
KW adult hippocampus; disease-associated SNP analysis; knockout mouse;
KW disease model mouse; cancer; neurological disorder.
XX
XX Homo sapiens.
OS
XX
XX WO200299103-A1.
PN
XX
XX 12-DEC-2002.
PD
XX
XX 27-MAY-2002; 2002WO-JP005134.
PF
XX
XX 04-JUN-2001; 2001JP-00168370.
PR
XX
XX 16-AUG-2001; 2001JP-00246915.
PR
XX
XX (KAZU-) KAZUSA DNA RES INST FOUND.
PA (PROT-) PROTEIN EXPRESS CO LTD.
XX
XX Ohara O, Nagase T, Nakajima D;
XX
XX WPI; 2003-140622/13.
DR N-PSDB; AAL51207.
XX
XX DNA preferentially expressed in human adult and fetal brain tissue useful
PT for diagnosis, treatment and analysis of cancer and mental disorders.
PT
XX
XX Claim 1; Page 56-60; 73pp; Japanese.
PS
XX
XX The invention comprises the amino acid and coding sequences of seven
CC human proteins that are preferentially expressed in adult whole brain,
CC foetal whole brain, tonsil and adult hippocampus tissue. The DNA
CC sequences are useful for the analysis of disease-associated single
CC nucleotide polymorphisms and the production of knockout and human disease
CC model mice. The DNA and protein sequences of the invention are useful for
CC the prevention (vaccine) and treatment of cancer and neurological
CC disorders. The present amino acid sequence represents a human protein of
CC the invention
XX
XX Sequence 452 AA;
SQ
Query Match 43.6%; Score 434; DB 6; Length 452;
Best Local Similarity 44.2%; Pred. No. 1.3e-36;
Matches 88; Conservative 39; Mismatches 54; Indels 18; Gaps 3;
QY 2 LALLEDCWCRIMSDVEQKSLMVTGIPADFEAEIOEVLQETLKSIGRYRLGKIFRKQENA 61
Db 7 LTLLEDWCRCMDVNSQRTLLVWGIPVNCDEAEIETLQAAMPQVS-YRMLGRMFWRENA 66
QY 62 NAVILLELLETDVSAIPSEVQGGKGVWVKVIFKTPNQDTEFLERINLFLEKEGQTVSGMFR 121
Db 67 KAVFIELADTVNTTLPSPHIFPGKGSWEVVVVKPRNPDEFLSRNLNYFLKDEGRSMTDVAR 126
QY 122 ALG-----QBALSPATVPCI-SPELLAHLIGQAMAHAPQELL-PMRYKLRVFGSGAVPAP 175
Db 127 ALGCCSLPASELDAEVMPPQVRSPL-----EPPKESMWYKRLKLVFGSGTAGSPSP 174
QY 176 EERSFEVMLEQATEIVKKEW 194
Db 175 GERTFEDWLEQVTEIMPIW 193
RESULT 14
ADA54410
ID ADA54410 standard; protein; 399 AA.
XX
AC ADA54410;
XX
XX 20-NOV-2003 (first entry)
DT
DE Human protein, SEQ ID 1978.
```

XX Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;
 KW Gene Therapy; human; secretory protein; membrane proteins; cancer;
 KW inflammatory disease; osteoporosis; neurological disease.
 XX
 OS Homo sapiens.
 XX
 PN EP1293569-A2.
 XX
 PD 19-MAR-2003.
 XX
 PF 21-MAR-2002; 2002EP-00006586.
 XX
 PR 14-SEP-2001; 2001JP-00328381.
 PR 24-JAN-2002; 2002US-0350435P.
 XX
 PA (HELI-) HELIX RES INST.
 PA (REAS-) RES ASSOC BIOTECHNOLOGY.
 XX
 PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
 PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
 PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;
 XX
 DR WPI; 2003-395539/38.
 DR N-PSDB; ADA52771.
 XX
 PT New polynucleotides encoding full-length polypeptides, e.g. secretory
 PT and/or membrane proteins, useful for developing medicines for diseases in
 PT which the gene is involved, or as target molecules for gene therapy.
 XX
 PS Claim 14; SEQ ID NO 1978; 205pp; English.
 XX
 CC The present invention relates to novel human secretory or membrane
 CC proteins (ADA54072-ADA55710) and their coding sequences (ADA52433-
 CC ADA54071). The coding sequences are useful in the gene therapy of
 CC diseases caused by abnormalities of the proteins, e.g. cancer,
 CC inflammatory diseases, osteoporosis or neurological disease.
 XX
 SQ Sequence 399 AA;
 Query Match 34.6%; Score 345; DB 6; Length 399;
 Best Local Similarity 38.9%; Pred. No. 2.6e-27;
 Matches 77; Conservative 38; Mismatches 63; Indels 20; Gaps 5;
 QY 2 LALLDWCIRMSVDQKSLMTVGTIPADFEAEIQVLOETLSGLRYRLGKIPKQENA 61
 DB 3 VTMLQDWCNMGVNRARGLLIGIPEDCDDAEFQESLEAALRPMGHFTVLGKAFREEDNA 62
 QY 62 NAVLELLEDTDVSAIPSEVQKGGVWKVIFKTPNQDTEF--LERINLFLEKEGQTVSGM 119
 DB 63 TAALVELDREVNIALVPREIPGTGPMNVFVPRCSGEFGLGRVTFHPEEGQMVESV 122
 QY 120 FRALGOEALSPATVPCISPPELLAHL--LQGMAMHAPQLL-PMRYRKURVFGSGSAVPAPE 176
 DB 123 AGALG-----VGLRRVCWLRISGQAV---QPWVEAVRCQSLGLVFGSGRDQPAPG 167
 QY 177 EESFEVWLEQATEIVKEW 194
 DB 168 EESFEVWLDHTTEMLHW 185
 RESULT 15
 ABG99947
 ID ABG99947 standard; protein; 399 AA.
 AC ABG99947;
 XX
 DT 17-JAN-2003 (first entry)
 XX
 DE Human novel polypeptide #60.
 XX
 KW Human; genetic disorder; gene mapping; medical imaging; cancer;
 KW neurodegenerative disorder; lymphoid cell disorder; osteoporosis;

KW Parkinson's disease; Alzheimer's disease; bone degenerative disorder;
 KW osteoarthritis; periodontal disease; liver fibrosis; viral infection;
 KW fungal infection; bacterial infection; autoimmune disease; diabetes;
 KW atopic dermatitis.
 XX
 OS Homo sapiens.
 XX
 PN WO200274961-A1.
 XX
 PD 26-SEP-2002.
 XX
 PF 14-MAR-2002; 2002WO-US005109.
 XX
 PR 15-MAR-2001; 2001US-00810173.
 PR
 XX (HYSE-) HYSEQ INC.
 PA
 XX Tang YT, Zhou P, Goodrich R, Asundi V, Zhang J, Zhao QA, Ren F;
 PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;
 PI Wehrman T, Wang J, Wang D, Drmanac RT;
 XX
 DR WPI; 2003-040556/03.
 DR N-PSDB; ABX05045.
 XX
 PT New isolated polypeptides and polynucleotides, useful for preventing,
 PT treating or ameliorating medical conditions, such as cancer,
 PT neurodegenerative disorders, lymphoid cell disorders, bone degenerative
 PT disorders, and infections.
 XX
 PS Claim 9; SEQ ID NO 586; 235pp; English.
 XX
 CC The invention relates to human polynucleotides and the polypeptides they
 CC encode. The polynucleotides and polypeptides are useful in diagnostics,
 CC forensics, gene mapping, medical imaging, identification of mutations,
 CC responsible for genetic disorders or other traits, assessing biodiversity
 CC and producing many other types of data and products dependent on DNA and
 CC amino acid sequences. They are also useful for preventing, treating or
 CC ameliorating medical conditions, such as cancer, neurodegenerative
 CC disorders (e.g. Parkinson's disease, Alzheimer's disease), lymphoid cell
 CC disorders, osteoporosis, osteoarthritis, bone degenerative disorders,
 CC periodontal disease, liver fibrosis, infections (e.g. viral, fungal or
 CC bacterial) or autoimmune diseases (e.g. diabetes, atopic dermatitis).
 CC Sequences ABG99888-ABG99989 and ABU00010-ABU00433 represent human
 CC polypeptides of the invention. Note: The sequence data for this patent is
 CC not represented in the printed specification but is based on sequence
 CC information supplied by the European Patent Office
 XX
 SQ Sequence 399 AA;
 Query Match 34.6%; Score 345; DB 6; Length 399;
 Best Local Similarity 38.9%; Pred. No. 2.6e-27;
 Matches 77; Conservative 38; Mismatches 63; Indels 20; Gaps 5;
 QY 2 LALLDWCIRMSVDQKSLMTVGTIPADFEAEIQVLOETLSGLRYRLGKIPKQENA 61
 DB 3 VTMLQDWCNMGVNRARGLLIGIPEDCDDAEFQESLEAALRPMGHFTVLGKAFREEDNA 62
 QY 62 NAVLELLEDTDVSAIPSEVQKGGVWKVIFKTPNQDTEF--LERINLFLEKEGQTVSGM 119
 DB 63 TAALVELDREVNIALVPREIPGTGPMNVFVPRCSGEFGLGRVTFHPEEGQMVESV 122
 QY 120 FRALGOEALSPATVPCISPPELLAHL--LQGMAMHAPQLL-PMRYRKURVFGSGSAVPAPE 176
 DB 123 AGALG-----VGLRRVCWLRISGQAV---QPWVEAVRCQSLGLVFGSGRDQPAPG 167
 QY 177 EESFEVWLEQATEIVKEW 194
 DB 168 EESFEVWLDHTTEMLHW 185
 Search completed: September 21, 2004, 13:35:13
 Job time : 37.0037 secs

B/ANK

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 13:20:22 ; Search time 45.1859 Seconds
(without alignments)
1361.621 Million cell updates/sec

Title: US-10-037-860-7

Perfect score: 996

Sequence: 1 PLALLEDWCRMSVDEQSKL.....EEESFEVWLEQATEIVKEMP 195

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriaph:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	982	98.6	364	4	O94959
2	972	97.6	194	4	O95145
3	959	96.3	364	6	Q9GMU3
4	766	76.9	365	11	Q8BHK0
5	593	59.5	283	4	Q9UL42
6	468	47.0	353	11	Q8VHZ4
7	466	46.8	353	4	Q8NG07
8	465	46.7	353	4	Q95144
9	462.5	46.4	455	4	Q9H0A4
10	462.5	46.4	463	4	Q9UL41
11	462	46.4	353	11	Q9CYP2
12	459.5	46.1	466	11	Q8JZW8
13	459	46.1	353	11	Q8CIC8
14	436	43.8	351	4	Q96BY2
15	436	43.8	351	4	Q9HAS1
16	434	43.6	448	4	Q8NET3

17	434	43.6	452	4	Q96PV4
18	431	43.3	351	6	Q95KI4
19	431	43.3	352	11	Q9ERH6
20	345	34.6	399	4	Q96AA0
21	338.5	34.0	246	11	Q8C533
22	338.5	34.0	430	11	Q80VM8
23	337	33.8	378	4	Q8N3H4
24	337	33.8	435	4	Q86V59
25	211.5	21.2	327	11	Q8VC32
26	146	14.7	538	4	Q9H0W5
27	146	14.7	538	4	Q8TB26
28	142	14.3	237	4	Q9H833
29	107.5	10.8	402	11	Q9CZA5
30	102	10.2	608	2	Q68043
31	100.5	10.1	393	11	Q9DB17
32	100.5	10.1	393	11	Q8VD24
33	99.5	10.0	403	4	Q8TE36
34	97.5	9.8	402	4	Q8NIC1
35	95	9.5	341	16	Q9WV83
36	92	9.2	192	4	Q8ND90
37	91	9.1	455	16	Q7V8D5
38	87.5	8.8	229	11	Q8RIJ2
39	87.5	8.8	269	16	Q88LH9
40	87.5	8.8	311	11	Q8VE43
41	87.5	8.8	897	10	Q7X6G9
42	87.5	8.8	1684	4	Q75218
43	87.5	8.8	3830	4	Q9Y6H4
44	87.5	8.8	3859	4	Q9Y631
45	86.5	8.7	959	16	Q7USU6

ALIGNMENTS

RESULT 1

O94959 PRELIMINARY; PRT; 364 AA.

AC O94959; DT 01-MAY-1999 (TREMBlrel. 10, Created)

DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)

DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)

DE Hypothetical protein KIAA0883 (Paraneoplastic associated brain-testis-cancer antigen).

GN KIAA0883.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Brain;

RX MEDLINE=99156230; PubMed=10048485;

RA Nagase T., Ishikawa K., Suyama M., Kikuno R., Hirotsawa M., Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;

RT "Prediction of the coding sequences of unidentified human genes. XII. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro."

RL DNA Res. 5:355-364 (1998).

RN [2]

SEQUENCE FROM N.A.

RA Eichen J.G., Dalmay J., Wade D., Rosenfeld M.R.;

RT "Characterization of a Brain-Testis-Cancer Antigen."

RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF286487; BAA74906.1; -.

DR EMBL; AF286487; BAA74906.1; -.

DR Genew; HGNC:9159; PNM2.

DR InterPro: IPR005162; Retrotrans_gag.

DR Pfam; PF03732; Retrotrans_gag; I.

KW Hypothetical protein.

SQ SEQUENCE 364 AA; 41509 MW; 6E417AD96B3F0E93 CRC64;

Query Match 98.6%; Score 982; DB 4; Length 364;
Best Local Similarity 99.0%; Pred. No. 7.1e-81;

```
Matches 192; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Qy 2 LALLEDWCRIMSDVDEQKSLMTGTPADPFEAEIOEVLQETLKSIGRYRLGKIFRKQENA 61
Db 3 LALLEDWCRIMSDVDEQKSLMTGTPADPFEAEIOEVLQETLKSIGRYRLGKIFRKQENA 62
Qy 62 NAVLELLEDTDVSAIPSEVQKGVWVKVIFKTPNQDTEFLERLNLFLKEGQTVSGMFR 121
Db 63 NAVLELLEDTDVSAIPSEVQKGVWVKVIFKTPNQDTEFLERLNLFLKEGQTVSGMFR 122
Qy 122 ALGQEALSPATVPCISPELLAHLGQAMAHAPQPLLPMRYKLRVFGSGAVPAPEESFE 181
Db 123 ALGQEGVSPATVPCISPELLAHLGQAMAHAPQPLLPMRYKLRVFGSGAVPAPEESFE 182
Qy 182 VWLEQATEIVKEWP 195
Db 183 VWLEQATEIVKEWP 196

RESULT 2
O95145 PRELIMINARY; PRT; 194 AA.
AC O95145;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE Paraneoplastic neuronal antigen NM2 (Fragment).
GN NM2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP TISSUE=Cerebellum;
RC Dalmat J., Rosenfeld M.R., Voltz R., Hoard R.;
RT "antigen recognized by serum of patients with paraneoplastic
neurologic disorders.";
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF037365; AAD02098.1; -.
FT NON-TER
SQ SEQUENCE 194 AA; 21808 MW; 285E4AD5BED3F28C CRC64;

Query Match 97.6%; Score 972; DB 4; Length 194;
Best Local Similarity 97.9%; Pred. No. 2.6e-80;
Matches 190; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
Qy 2 LALLEDWCRIMSDVDEQKSLMTGTPADPFEAEIOEVLQETLKSIGRYRLGKIFRKQENA 61
Db 1 LALLEDWCRIMSDVDEQKSLMTGTPADPFEAEIOEVLQETLKSIGRYRLGKIFRKQENA 60
Qy 62 NAVLELLEDTDVSAIPSEVQKGVWVKVIFKTPNQDTEFLERLNLFLKEGQTVSGMFR 121
Db 61 NAVLELLEDTDVSAIPSEVQKGVWVKVIFKTPNQDTEFLERLNLFLKEGQTVSGMFR 120
Qy 122 ALGQEALSPATVPCISPELLAHLGQAMAHAPQPLLPMRYKLRVFGSGAVPAPEESFE 181
Db 121 ALGQEALSPATVPCISPELLAHLGQAMAHAPQPLLPMRYKLRVFGSGAVPAPEESFE 180
Qy 182 VWLEQATEIVKEWP 195
Db 181 VWLEQATEIVKEWP 194

RESULT 3
Q9GMU3 PRELIMINARY; PRT; 364 AA.
AC Q9GMU3;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
```

```
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
OC Cercopitheidae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
RA Suzuki Y., Sugano S., Hashimoto K.;
RT "Isolation of full-length cDNA clones from macaque brain cDNA
libraries.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047632; BAB12156.1; -.
DR InterPro; IPR005162; Retrotrans_gag.
DR Pfam; PF03732; Retrotrans_gag; 1.
DR Hypothetical protein.
SQ SEQUENCE 364 AA; 41350 MW; 0CF72210D7EC1524 CRC64;

Query Match 96.3%; Score 959; DB 6; Length 364;
Best Local Similarity 95.9%; Pred. No. 8.8e-79;
Matches 186; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
Qy 2 LALLEDWCRIMSDVDEQKSLMTGTPADPFEAEIOEVLQETLKSIGRYRLGKIFRKQENA 61
Db Osada N. :
Db 3 LALLEDWCRIMSDVDEQKSLMTGTPADPFEAEIOEVLQETLKSIGRYRLGKIFRKQENA 62
Qy 62 NAVLELLEDTDVSAIPSEVQKGVWVKVIFKTPNQDTEFLERLNLFLKEGQTVSGMFR 121
Db 63 NAVLELLEDTDVSAIPSEVQKGVWVKVIFKTPNQDTEFLERLNLFLKEGQTVSGMFR 122
Qy 122 ALGQEALSPATVPCISPELLAHLGQAMAHAPQPLLPMRYKLRVFGSGAVPAPEESFE 181
Db 123 ALGHEGMSPATVPCISPELLAHLGQAMAHAPQPLLPMRYKLRVFGSGAVPAPEESFE 182
Qy 182 VWLEQATEIVKEWP 195
Db 183 VWLEQATEIVKEWP 196

RESULT 4
Q8BHK0 PRELIMINARY; PRT; 365 AA.
AC Q8BHK0;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similar to KIAA0883 protein.
GN PNM2 OR A830049P17RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cortex;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL; AK043718; BAC31626.1; -.
DR EMBL; AK043910; BAC31700.1; -.
DR MGD; MGI:2444129; Pnm2.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005215; P:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000425; MIP.
DR InterPro; IPR005162; Retrotrans_gag.
DR Pfam; PF03732; Retrotrans_gag; 1.
DR PROSITE; PS00221; MIP; 1.
SQ SEQUENCE 365 AA; 41201 MW; 0F90C940B9D843D9 CRC64;
```

```

Query Match      76.9%; Score 766; DB 11; Length 365;
Best Local Similarity 74.2%; Pred. No. 3.1e-61;
Matches 144; Conservative 20; Mismatches 30; Indels 0; Gaps 0;

QY 2 LALLEDWCRIMSDVEOKSLMVTGIPADFEAEIQTVEVLQETLKSGLGRYLLGKIFRKQENA 61
   :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 3 VALLEWCKIMGVDVQKSLLVDPVDCGPEIQTVLQEAALKCVGSYRLGKIFQKQDNT 62
   :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

QY 62 NAVLELLEDTDVSAIPSEVQGGVGVKVIKFTPNQDTEFLERLNLFLKEGQTVSGMFR 121
   :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 63 SVVLVELMEDTMSVVPSEVQGGVGVKVIKFTPNQDTEFLERLNLFLKEGQTVAGMFR 122
   :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

QY 122 ALGOEALSPATVPCISPELLAHLGQAMAHAPQLLPMRYRKLRFVSGSAVPAPEESFE 181
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 123 ALKHGVSATPPCTSPPELLAHLTGQAMVHGQRPPLPVKYCKWRIFGSGTAAAPPEEPFE 182
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

QY 182 VMLQEQATEIVKWP 195
   |||||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 183 VMLQEQATEIAKWP 196
   |||||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

RESULT 5
Q9UL42 PRELIMINARY; PRT; 283 AA.
AC Q9UL42;
DT 01-MAY-2000 (TremBLrel. 13, Created)
DT 01-MAY-2000 (TremBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TremBLrel. 21, Last annotation update)
DE Paraneoplastic cancer-testis-brain antigen (Fragment).
GN MA4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Gultekin S.H., Voltz R., Rosenfeld M.R., Gerstner E., Eichen J.,
RA Posner J.B., Dalmay J.;
RT "Identification of a novel cancer testis brain antigen using serum
RT antibodies from patients with testicular tumors and paraneoplastic
RT limbic encephalomyelitis";
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF083115; AAF05626.1; -.
DR Rattus norvegicus (Rat).
DR Pfam; PF03732; Retrotrans_gag; 1.
FT NON TER
FT 1
SQ SEQUENCE 283 AA; 32333 MW; E27BA6BCDCD240A4 CRC64;

Query Match      59.5%; Score 593; DB 4; Length 283;
Best Local Similarity 98.3%; Pred. No. 1.2e-45;
Matches 113; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 81 VQGGGVKWKVIFKFTPNQDTEFLERLNLFLKEGQTVSGMFRALGOEALSPATVPCISPEL 140
   |||||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1 VQGGGVKWKVIFKFTPNQDTEFLERLNLFLKEGQTVSGMFRALGOEALSPATVPCISPEL 60
   |||||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

QY 141 LAHLGQAMAHAPQLLPMRYRKLRFVSGSAVPAPEESFEVWLQEQATEIVKWP 195
   |||||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 61 LAHLGQAMAHAPQLLPMRYRKLRFVSGSAVPAPEESFEVWLQEQATEIVKWP 115
   |||||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

RESULT 6
Q8VHZ4 PRELIMINARY; PRT; 353 AA.
AC Q8VHZ4;
DT 01-MAR-2002 (TremBLrel. 20, Created)
DT 01-MAR-2002 (TremBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TremBLrel. 20, Last annotation update)
DE Paraneoplastic onconeural protein MA1.
GN MA1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

```

```

OX NCBI_TaxID=101116;
RN [1]
RP SEQUENCE FROM N.A.
RA Peilkofer H.L., Schubart A., Pagany M., Voltz R.D., Linington C.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF335505; AAL73196.1; -.
SQ SEQUENCE 353 AA; 39793 MW; FFC3A717FDCEADA3 CRC64;

Query Match      47.0%; Score 468; DB 11; Length 353;
Best Local Similarity 47.9%; Pred. No. 3.5e-34;
Matches 93; Conservative 38; Mismatches 57; Indels 6; Gaps 4;

QY 2 LALLEDWCRIMSDVEOKSLMVTGIPADFEAEIQTVEVLQETLKSGLGRYLLGKIFRKQENA 61
   :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 3 MTLLEDWCRGMDVNSQRALLVWGIPVNCDEAEIETTLQAAMPQVS-YRMLGRMFWEENA 61
   :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

QY 62 NAVLELLEDTDVSAIPSEVQGGVGVKVIKFTPNQDTEFLERLNLFLKEGQTVSGMFR 121
   :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 62 KAALLELTGTVDYSLIPREMPGKGLWKVVPKPTSDAEFLERLHLFLAREGTVQDVAR 121
   :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

QY 122 ALGOEALSPATVPCISPELLAHLGQAMAHAPQLL-PMRYRKLRFVSGSAVPAPEESF 180
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 122 VLGFQ--NPAPAP--GPMPAEMLNLYLDNVIQPLVESIWKYKLTFLSGRDIPOGGETF 177
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

QY 181 EYVLEQATEIVKWP 194
   |||||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 178 DSWLEHSNEVIEEW 191
   |||||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

RESULT 7
Q8NG07 PRELIMINARY; PRT; 353 AA.
AC Q8NG07;
DT 01-OCT-2002 (TremBLrel. 22, Created)
DT 01-OCT-2002 (TremBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TremBLrel. 23, Last annotation update)
DE Paraneoplastic antigen.
GN PNMA1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Schueller M.M., Jenne D.E., Schutze dit Belkner N., Hohlfeld R.,
RA Voltz R.;
RT "Mal (PNMA1).";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Prostate;
RA Strausberg R.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF320308; AAN05100.1; -.
DR EMBL; BC039577; AAH39577.1; -.
SQ SEQUENCE 353 AA; 39761 MW; EB7F5B6AEDA25961 CRC64;

Query Match      46.8%; Score 466; DB 4; Length 353;
Best Local Similarity 47.9%; Pred. No. 5.4e-34;
Matches 93; Conservative 38; Mismatches 57; Indels 6; Gaps 4;

QY 2 LALLEDWCRIMSDVEOKSLMVTGIPADFEAEIQTVEVLQETLKSGLGRYLLGKIFRKQENA 61
   :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 3 MTLLEDWCRGMDVNSQRALLVWGIPVNCDEAEIETTLQAAMPQVS-YRMLGRMFWEENA 61
   :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

QY 62 NAVLELLEDTDVSAIPSEVQGGVGVKVIKFTPNQDTEFLERLNLFLKEGQTVSGMFR 121
   :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 62 KAALLELTGTVDYSLIPREMPGKGLWKVVPKPTSDAEFLERLHLFLAREGTVQDVAR 121
   :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

QY 122 ALGOEALSPATVPCISPELLAHLGQAMAHAPQLL-PMRYRKLRFVSGSAVPAPEESF 180
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 122 VLGFQ--NPTTP--GPMPAEMLNLYLDNVIQPLVESIWKYKLTFLSGRDIPOGGETF 177
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

```

US Homo sapiens (human) .
 OC Eukaryota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Eukaryota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;


```
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Brain;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL; AK028331; BAC25885.1; -.
DR MGD; MGI:2180564; Pnuma;
SQ SEQUENCE 353 AA; 39688 MW; 18CEDC3AC4E70939 CRC64;

Query Match 46.1%; Score 459; DB 11; Length 353;
Best Local Similarity 47.9%; Pred. No. 2.3e-33;
Matches 93; Conservative 36; Mismatches 59; Indels 6; Gaps 4;

QY 2 LALLEDCWCRIMSVDEQKSLMTGIPADFEAEIQTQVLQETLKSIGRYLLGKIFRKQENA 61
   :||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 3 MTLLEDWCRGMDVNSORALLVWGPVNCDETEIETLQAAMPQVS-YRVILGRMFWRENA 61
   :||||| : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 62 NAVLELLEDTDVSAIPSEVQGGVWVKVFKTPNQDTEFLERINLFLKEGQTVSGMFR 121
   :||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 62 KAALLELTGAVDSLIPREMPGKGLWVKVFKPTSDAVFLERHLFLAREGTVQDVAR 121
   :||||| : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 122 ALGQEALSPATVPCISPELLAHLGQAMAHAPQPLP-PMYRKLKRVFSGSAVPAPPEESF 180
   :||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 122 VLGFQ--NPAPAP--GPETPAEMLNYLDNVQPLVESIWYKKLTLPFGKDIPGGETF 177
   :||||| : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 181 EVMLEQATEIVKQW 194
   :||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 178 DSWLEHSEIIEEW 191
   :||||| : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 14
Q96BY2 PRELIMINARY; PRT; 351 AA.
ID Q96BY2
AC Q96BY2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Modulator of apoptosis 1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RA Strausberg R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; BC015044; AAH15044.1; -.
DR Genew; HGNC:16658; MOAP1.
SQ SEQUENCE 351 AA; 39512 MW; 5310142AC02B563C CRC64;

Query Match 43.8%; Score 436; DB 4; Length 351;
Best Local Similarity 46.9%; Pred. No. 2.9e-31;
Matches 91; Conservative 30; Mismatches 71; Indels 2; Gaps 2;

QY 2 LALLEDCWCRIMSVDEQKSLMTGIPADFEAEIQTQVLQETLKSIGRYLLGKIFRKQENA 61
   :||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 3 LRLLEDWCRGMDMNPRAKLLIAGISQSCSVAEIEEALQAGLAPLGEYRLIGRMFRDNR 62
   :||||| : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 62 NAVLELLEDTDVSAIPSEVQGGVWVKVFKTPNQDTEFLERINLFLKEGQTVSGMFR 121
   :||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 63 KVALVGLTAETSHALVPKEIPKGGIWRVIFKPPDPDNTFLSRLNEFLAGEGTMVGLSR 122
   :||||| : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 122 ALGQEALSPATVPCISPELLAHLGQAMAHAPQPLP-MRYRKLKRVFSGSAVPAPPEESF 180
   :||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 123 ALGHENGLDPEQGMIPENWAPMLAQAL-EALQPALQCLKYKKLRVFSGRESPEPGEF 181
   :||||| : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 181 EVMLEQATEIVKQW 194
   :||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 182 GRWMFHTTOMIKAW 195
   :||||| : : : : : : : : : : : : : : : : : : : : : : : : : :

Search completed: September 21, 2004, 13:39:31
Job time : 46.1859 secs
```

```
Db 123 ALGHENGLDPEQGMIPENWAPMLAQAL-EALQPALQCLKYKKLRVFSGRESPEPGEF 181
QY 181 EVMLEQATEIVKQW 194
   :||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 182 GRWMFHTTOMIKAW 195
   :||||| : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 15
Q9HAS1 PRELIMINARY; PRT; 351 AA.
ID Q9HAS1
AC Q9HAS1;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE MAP-1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Cerebellum;
RX MEDLINE=21264738; PubMed=11060313;
RA Tan K.O., Tan K.M.L., Chan S.-L., Yee X.S.Y., Bevort M., Ang K.C.,
RA Yu V.C.;
RT "MAP-1, a novel proapoptotic protein containing a BH3-like motif that
RT associates with Bax through its Bcl-2 homology domains.";
RL J. Biol. Chem. 276:2802-2807 (2001).
DR EMBL; AF305550; AAG31786.1; -.
SQ SEQUENCE 351 AA; 39521 MW; 8736401AD8A24EAD CRC64;

Query Match 43.8%; Score 436; DB 4; Length 351;
Best Local Similarity 46.9%; Pred. No. 2.9e-31;
Matches 91; Conservative 30; Mismatches 71; Indels 2; Gaps 2;

QY 2 LALLEDCWCRIMSVDEQKSLMTGIPADFEAEIQTQVLQETLKSIGRYLLGKIFRKQENA 61
   :||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 3 LRLLEDWCRGMDMNPRAKLLIAGISQSCSVAEIEEALQAGLAPLGEYRLIGRMFRDNR 62
   :||||| : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 62 NAVLELLEDTDVSAIPSEVQGGVWVKVFKTPNQDTEFLERINLFLKEGQTVSGMFR 121
   :||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 63 KVALVGLTAETSHALVPKEIPKGGIWRVIFKPPDPDNTFLSRLNEFLAGEGTMVGLSR 122
   :||||| : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 122 ALGQEALSPATVPCISPELLAHLGQAMAHAPQPLP-MRYRKLKRVFSGSAVPAPPEESF 180
   :||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 123 ALGHENGLDPEQGMIPENWAPMLAQAL-EALQPALQCLKYKKLRVFSGRESPEPGEF 181
   :||||| : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 181 EVMLEQATEIVKQW 194
   :||||| : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 182 GRWMFHTTOMIKAW 195
   :||||| : : : : : : : : : : : : : : : : : : : : : : : : : :

Search completed: September 21, 2004, 13:39:31
Job time : 46.1859 secs
```